

**Rekombinante Herstellung und Charakterisierung  
flankierender Domänen der *M. galloprovincialis*  
preCol-Moleküle**

**DISSERTATION**

zur Erlangung des akademischen Grades eines  
Doktors der Naturwissenschaften (Dr. rer. nat.)  
an der Bayreuther Graduiertenschule  
für Mathematik und Naturwissenschaften (BayNAT)  
der Universität Bayreuth

vorgelegt von  
Markus Heim (M. Sc.)  
aus Kaufbeuren

Bayreuth, Februar 2014



Die vorliegende Arbeit wurde in der Zeit von März 2009 bis Februar 2014 in Bayreuth am Lehrstuhl Biomaterialien unter Betreuung von Herrn Professor Dr. Thomas Scheibel angefertigt.

Vollständiger Abdruck der von der Bayreuther Graduiertenschule für Mathematik und Naturwissenschaften (BayNAT) der Universität Bayreuth genehmigten Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.)

Dissertation eingereicht am: 19.02.2014

Zulassung durch das Leitungsgremium: 07.03.2014

Wissenschaftliches Kolloquium: 01.08.2014

Amtierender Direktor: Prof. Dr. Franz Xaver Schmid

Prüfungsausschuss:

Prof. Dr. Thomas Scheibel (Erstgutachter)

Prof. Dr. Franz Xaver Schmid (Zweitgutachter)

Prof. Dr. Birgitta Wöhrl (Vorsitz)

Prof. Dr. Andreas Fery



## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung.....</b>	<b>1</b>
1.1	Extrakorporale Proteinstrukturen – Seiden und seidenähnliche Proteine .....	1
1.1.1	Arthropodenseide .....	1
1.1.2	Seidenähnliche Proteinstrukturen bei Muscheln .....	5
1.2	Aufbau von Seiden und seidenähnlichen Strukturen .....	6
1.2.1	Makroskopischer Kern-Hülle-Aufbau .....	7
1.2.2	Molekularer Aufbau .....	10
1.2.3	Struktur-Funktions-Beziehungen bei Seiden und seidenähnlichen Proteinen.....	16
1.3	Der Muschelbyssus.....	22
1.3.1	Makroskopischer Aufbau des Byssusfadens.....	23
1.3.2	Die preCols als zentrale Proteinkomponente der Byssusfäden .....	24
1.3.3	Isolierung nativer und Herstellung rekombinanter preCols.....	28
1.4	Zielsetzung .....	29
<b>2</b>	<b>Material und Methoden.....</b>	<b>31</b>
2.1	Verwendete Materialien .....	31
2.1.1	Chemikalien .....	31
2.1.2	Größenstandards und Kits .....	33
2.1.3	Oligonukleotide und artifizielle Nukleotidsequenzen .....	33
2.1.4	Plasmide und Vektorkonstrukte .....	36
2.1.5	<i>E. coli</i> -Stämme .....	38
2.1.6	Proteine und Peptide.....	39
2.1.7	Chromatographiematerialien .....	41
2.1.8	Sonstige Materialien.....	41
2.1.9	Geräte .....	42
2.1.10	Computersoftware und Online-Tools.....	45
2.2	Bioinformatische Methoden .....	46
2.2.1	Bestimmung der physikalisch-chemischen Eigenschaften von Proteinen .....	46
2.2.2	Erstellen von DNA- und Aminosäuresequenzalignments.....	47
2.2.3	Sekundärstrukturvorhersage von Proteinen .....	47
2.2.4	Vergleich der Codonverwendung ( <i>codon usage</i> ) .....	47
2.3	Molekularbiologische Methoden .....	48
2.3.1	Verwendete Lösungen und Puffer.....	48
2.3.2	Auftrennung von DNA .....	48
2.3.3	Reinigung und Isolierung von DNA.....	49
2.3.4	Polymerasekettenreaktion (PCR) .....	50
2.3.5	DNA-Restriktionsverdau .....	54
2.3.6	Ligation von DNA-Enden.....	55
2.3.7	Sequenzierung von DNA.....	56
2.4	Mikrobiologische Methoden .....	56
2.4.1	Verwendete Lösungen und Puffer.....	56
2.4.2	Kultivierung und Aufbewahrung von <i>E. coli</i> -Zellen .....	57

2.4.3	Herstellung transformationskompetenter <i>E. coli</i> -Zellen .....	58
2.4.4	Transformation von <i>E. coli</i> -Zellen .....	59
<b>2.5</b>	<b>Rekombinante Produktion von <i>flank</i>-Proteinen .....</b>	<b>60</b>
2.5.1	Verwendete Lösungen und Puffer .....	60
2.5.2	Testexpression .....	61
2.5.3	Fermentation transformierter <i>E. coli</i> -Stämme .....	62
2.5.4	Ernten und Waschen von <i>E. coli</i> -Zellen.....	63
2.5.5	Aufschluss von <i>E. coli</i> -Zellen .....	64
<b>2.6</b>	<b>Reinigung rekombinant hergestellter Proteine .....</b>	<b>64</b>
2.6.1	Verwendete Lösungen und Puffer .....	65
2.6.2	Affinitätschromatographie.....	65
2.6.3	Gelpermeationschromatographie.....	67
2.6.4	Abspaltung terminaler Protein- <i>tags</i> .....	68
2.6.5	Reinigungsstrategie für cwCT(F34W).....	71
<b>2.7</b>	<b>Proteinanalytische Methoden.....</b>	<b>72</b>
2.7.1	Verwendete Lösungen und Puffer .....	72
2.7.2	SDS-Polyacrylamid-Gelelektrophorese .....	73
2.7.3	Silberfärbung.....	76
2.7.4	Westernblot .....	77
<b>2.8</b>	<b>Spektroskopische Methoden .....</b>	<b>79</b>
2.8.1	Verwendete Lösungen und Puffer .....	79
2.8.2	UV-Absorptionsspektroskopie .....	79
2.8.3	Zirkulardichroismusspektroskopie .....	81
2.8.4	Fluoreszenzspektroskopie.....	84
<b>2.9</b>	<b>MALDI-TOF Massenspektrometrie .....</b>	<b>85</b>
<b>2.10</b>	<b>Herstellung von SDS-Mizellen und Emulsionen kleiner unilamellarer Lipidvesikel (SUV) .....</b>	<b>86</b>
<b>2.11</b>	<b>Präparation von Polystyrol-Nanokügelchen .....</b>	<b>88</b>
<b>3</b>	<b>Ergebnisse.....</b>	<b>91</b>
<b>3.1</b>	<b>Bioinformatische Charakterisierung der flankierenden Regionen der preCols.....</b>	<b>91</b>
3.1.1	Vergleich der cDNA-Bank preCol-Sequenzen mit bekannten Sequenzen .....	91
3.1.2	Sekundärstrukturvorhersage der flankierenden Regionen der preCols.....	93
3.1.3	Sequenzanalyse der flankierenden Regionen bezüglich Aminosäurezusammensetzung und Codonverwendung.....	95
<b>3.2</b>	<b>Herstellung rekombinanter <i>M. galloprovincialis flank</i>-Proteine.....</b>	<b>98</b>
3.2.1	Herstellung unveränderter preCol <i>flank</i> -Proteine .....	100
3.2.2	Herstellung optimierter preCol <i>flank</i> -Proteine.....	101
<b>3.3</b>	<b>Reinigung rekombinant hergestellter SUMO-<i>flank</i>-Fusionsproteine .....</b>	<b>105</b>
<b>3.4</b>	<b>Produktion rekombinanter <i>flank</i>-Proteine durch Herstellung und Reinigung doppelt getagter Fusionsproteine.....</b>	<b>108</b>
<b>3.5</b>	<b>Struktur und Stabilität des <i>flank</i>-Proteins cwCT(F34W).....</b>	<b>111</b>

3.6	<b>Strukturelle Interaktion von cwCT(F34W) mit Mizellen und kleinen unilamellaren Lipidvesikeln (SUV) .....</b>	<b>114</b>
3.6.1	Interaktion von cwCT(F34W) mit SDS-Mizellen .....	114
3.6.2	Interaktion von cwCT(F34W) mit kleinen unilamellaren Lipidvesikeln (SUV) .....	117
3.6.3	Abhängigkeit der DSPG-SUV-induzierten cwCT(F34W)- Sekundärstrukturänderung von verschiedenen Parametern .....	120
3.7	<b>Partitioning-folding coupling von cwCT(F34W) an Lipidmembrangrenzflächen.....</b>	<b>124</b>
4	<b>Diskussion .....</b>	<b>127</b>
4.1	<b>Rekombinante Herstellung der flankierenden Domänen der M. galloprovincialis preCols .....</b>	<b>127</b>
4.1.1	Unterschiede zwischen den zur rekombinanten Herstellung der <i>flank</i> -Proteine verwendeten cDNA-Sequenzen und bereits bekannten Isoformen .....	127
4.1.2	Rekombinante Herstellung von <i>flank</i> -Proteinen in <i>E. coli</i> .....	128
4.1.3	Reinigung rekombinant hergestellter <i>flank</i> -Proteine.....	130
4.2	<b>Struktur und Stabilität des preCol-NG <i>flank</i>-Proteins cwCT(F34W) in Lösung.....</b>	<b>132</b>
4.3	<b>Interaktion von cwCT(F34W) mit Mizellen und kleinen unilamellaren Vesikeln (SUVs) .....</b>	<b>135</b>
4.3.1	Interaktion von cwCT(F34W) mit SDS-Mizellen .....	135
4.3.2	Interaktion von cwCT(F34W) mit kleinen unilamellaren Lipidvesikeln (SUV) .....	136
4.4	<b>Bedeutung des PFC-Verhaltens der carboxyterminal flankierenden Region für die Funktion von preCol-NG.....</b>	<b>143</b>
5	<b>Zusammenfassung .....</b>	<b>147</b>
6	<b>Summary .....</b>	<b>149</b>
7	<b>Literatur.....</b>	<b>151</b>
8	<b>Abkürzungen.....</b>	<b>165</b>
9	<b>Publikationen.....</b>	<b>169</b>
10	<b>Danksagung .....</b>	<b>171</b>
11	<b>Anhang .....</b>	<b>173</b>
11.1	<b>Verwendete Proteine .....</b>	<b>173</b>
11.1.1	Unveränderte <i>flank</i> -Proteine .....	173
11.1.2	Optimierte <i>flank</i> -Fusionsproteine .....	175
11.1.3	Fusions-tags.....	176
11.1.4	Verwendete Kontrollproteine .....	177
12	<b>Erklärung .....</b>	<b>179</b>





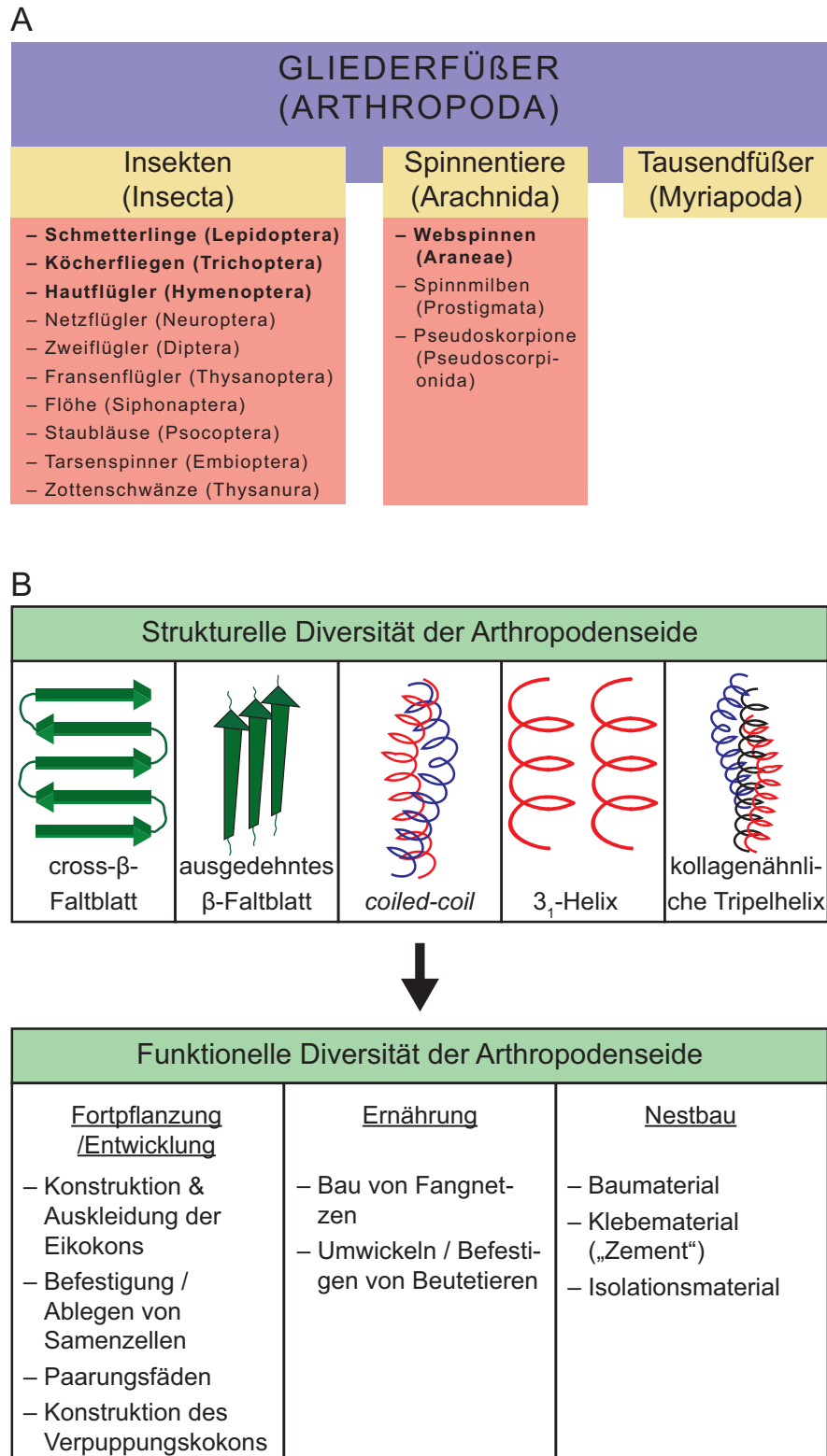
# 1 Einleitung

## 1.1 Extrakorporale Proteinstrukturen – Seiden und seidenähnliche Proteine

In seiner ursprünglichen Bedeutung bezog sich der Begriff Seide hauptsächlich auf die Proteinfasern, die von den Larven des Maulbeerspinners *Bombyx mori* produziert und zum Bau der Kokons verwendet werden, sowie auf die aus diesen Proteinfasern hergestellten Textilien. Ähnlich wie *B. mori* verwenden aber auch andere Vertreter der Arthropoden (Gliederfüßer) auf Proteinfasern beruhende, extrakorporale Strukturen, die ebenfalls der Erfüllung überlebenswichtiger Aufgaben dienen<sup>[1-5]</sup>. Aufgrund ihrer makroskopischen und molekularen Ähnlichkeiten zur *B. mori* Kokonseide wurde die ursprüngliche Definition des Begriffs Seide deswegen erweitert und inzwischen bezeichnet man mit diesem Oberbegriff allgemein fibrilläre, aus hochrepetitiven Aminosäuresequenzen aufgebaute Proteine, die hochkonzentriert in löslicher Form innerhalb eines Tieres gespeichert und dann mithilfe eines als Spinnen bezeichneten Prozesses in unlösliche Proteinfasern überführt werden<sup>[2]</sup>. Da sich diese Definition jedoch ausdrücklich auf die von Vertretern der Arthropoden gebildeten, extrakorporalen Proteinstrukturen sowie deren zugrunde liegende Proteine bezieht, wird zur Bezeichnung vergleichbarer Proteine und Proteinstrukturen der Lebewesen, die nicht dem Stamm der Arthropoden angehören, eher der Begriff „seidenähnlich“ verwendet. Aufgrund ihrer weiten Verbreitung und der diversen biologischen Aufgaben, die sie erfüllen, soll dieser Paragraph einen ersten Überblick über Herkunft, Verbreitung und Funktionsspektrum sowohl von Seiden wie auch seidenähnlichen Proteinstrukturen geben.

### 1.1.1 Arthropodenseide

Vertreter verschiedenster Ordnungen innerhalb der Klassen Insekten (*Insecta*), Spinnentiere (*Arachnida*) und Tausendfüßer (*Myriapoda*) produzieren und verwenden Seiden (vgl. Abb. 1-1 A).



**Abb. 1-1. Zusammenfassung seidenproduzierender Vertreter der Arthropoden und Übersicht über die strukturelle und funktionelle Diversität der Arthropodenseide. (A)** Alle seidenproduzierenden Individuen gehören zu Ordnungen (rot) innerhalb der Klassen (gelb) Insekten, Spinnentiere oder Tausendfüßer. **(B)** Verschiedene Arthropodenseidenproteine können innerhalb der Seidenfäden spezifische Konformationen einnehmen und der Seide so die für das jeweilige Aufgabenfeld benötigten Eigenschaften verleihen.

Die große Zahl verschiedener Arthropodenseidenproteine, die Molekulargewichte von 30 kDa bis zu über 350 kDa aufweisen<sup>[6-9]</sup>, ist das Ergebnis einer konvergenten Evolution. Trotzdem weisen die Proteine sowie die aus ihnen bestehenden Seiden einige charakteristische Gemeinsamkeiten auf. Alle Seidenproteine besitzen ungewöhnlich hohe Anteile der Aminosäuren Glycin (G), Alanin (A) und Serin (S). Die jeweiligen Seidenproteine werden hochkonzentriert als wässrige Lösungen in spezialisierten Drüsen gelagert. Durch Aneinanderlagern und Ausrichten der Proteine bilden sich flüssigkristalline Phasen sowie während des Spinnprozesses kristalline Bereiche innerhalb der Seidenfäden aus<sup>[10,11]</sup>. Für das Verspinnen der hochkonzentrierten und oft flüssigkristallinen Proteinlösungen zu Seidenfäden sind Scherkräfte nötig, um über die scherverdünnend (pseudoplastisch) wirkende Ausrichtung der Moleküle ein Herabsetzen der Viskosität der Proteinlösung zu erreichen<sup>[12,13]</sup>. Innerhalb der fertigen Seidenfäden herrschen dann ein oder mehrere dominante Sekundärstrukturelemente bzw. aus diesen entstandene Quartärstrukturen vor, die gleichzeitig charakteristisch für den jeweiligen Seidentyp sind. Bisher nachgewiesene Proteinkonformationen in Arthropodenseiden sind parallel oder senkrecht zur Fadenachse verlaufende  $\beta$ -Faltblätter,  $\alpha$ -helikale *coiled-coils*,  $3_1$ -Helices, kollagenähnliche tripelhelikale Strukturen<sup>[5]</sup> (vgl. Abb. 1-1 B) sowie Abfolgen von  $\beta$ -Schleifen<sup>[14]</sup>. Die fertig gesponnenen Fäden haben schließlich, abhängig vom Seidentyp, Durchmesser von ca. 2  $\mu\text{m}$  bis 200  $\mu\text{m}$  und werden von Vertretern der Arthropoden zur Bewältigung verschiedenster überlebenswichtiger Aufgaben genutzt<sup>[2,5]</sup> (vgl. Abb. 1-1 B).

#### 1.1.1.1 Insektenseiden

Zu den am umfangreichsten und besten untersuchten Insektenseiden zählen diejenigen der Schmetterlinge (*Lepidoptera*), der Köcherfliegen (*Trichoptera*) und der Hautflügler (*Hymenoptera*). Die Larven fast aller Vertreter dieser Ordnungen produzieren, zumindest während eines bestimmten Zeitraums ihres Larvenstadiums, Seiden<sup>[3]</sup>. Das Anwendungsspektrum der von den Larven hergestellten Seiden ist dabei recht breit gefächert (vgl. Abb. 1-1 B). Hauptsächlich benutzen Larven die produzierte Seide, um sich in einen Kokon, der als geschützter Raum während der Verpuppung und Puppenruhe dient, einzuspinnen, oder um ähnliche Schutzbehausungen und Tunnel zu bauen. Die bekannteste Kokonseide ist sicherlich die der Raupen des Seidenspinners *Bombyx mori* (einem

Vertreter der Schmetterlinge), da sie von Menschen bereits seit der Antike als Rohstoff für edle Stoffe und Textilien verwendet wird. Aufgrund ihrer langen kulturhistorischen Geschichte und ihrer enormen wirtschaftlichen Bedeutung ist diese Seide zugleich auch die am umfangreichsten erforschte Arthropodenseide überhaupt.

Die Seiden der Insektenlarven übernehmen neben ihrer reinen Schutzfunktion auch wichtige Rollen bei der Ernährung und beim Nestbau. Die Larven von Köcherfliegen nutzen ihre Seide zum Bau von Fangnetzen und als Verbindungsmaterial, um organische und anorganische Substanzen (Blätter, Sandkörner, Steine, etc.) zu Schutzbehausungen und Verstecken zusammenzufügen<sup>[3]</sup>. Honigbienen- und Hornissenlarven (beides Vertreter der Hautflügler) kleiden ihre aus Wachs geformten Brutwaben zusätzlich mit Seiden aus. Neben der dadurch erreichten mechanischen Stabilisierung der Brutstätten<sup>[15,16]</sup> übernimmt die Seide auch thermoregulatorische Eigenschaften<sup>[15-18]</sup> und sorgt für eine konstant (hohe) Luftfeuchtigkeit<sup>[4,19]</sup>. Beides ist für die korrekte Entwicklung und das Überleben der Larven unabdingbar.

Die Fähigkeit, Seide über das Larvenstadium hinaus zu produzieren, ist hingegen weit weniger verbreitet und bleibt nur bei einigen Spezies erhalten. Adulte, weibliche Wespen (Hautflügler) produzieren z.B. noch Seide, um diese, wie die Larven auch, zum Verbinden von pflanzlichen Fasern während des Nestbaus sowie zur Nestbefestigung zu verwenden<sup>[4,5]</sup>.

### 1.1.1.2 *Spinnenseiden*

Webspinnen (eine Ordnung der Spinnentiere, vgl. Abb. 1-1 A) produzieren und verwenden ebenso wie zahlreiche Vertreter der Insekten Seide. Im Gegensatz zu letzteren produzieren Webspinnen jedoch nicht nur eine einzelne, sondern zwischen zwei und sieben verschiedene Seidenarten, für deren Produktion jeweils auch unterschiedliche Drüsen verwendet werden<sup>[20]</sup>. Jeder Seidentyp übernimmt, alleine oder im Verbund mit weiteren Seidentypen, eine spezifische Funktion. Zwei universell bei Webspinnen anzutreffende Anwendungen der Seide sind die Herstellung von Eikokons, um die Eier vor äußeren Einflüssen zu schützen, sowie das Umwickeln gefangener Beutetiere. Letzteres dient der Fixierung der Beute bis zum Wirkungseintritt des injizierten Verdauungssafts. Über diese beiden weit verbreiteten Anwendungen hinausgehend werden Seiden noch zur

Konstruktion komplexer Beutefangvorrichtungen verwendet. Vertreter aus der Überfamilie der Radnetzspinnen (*Araneoidae*) benutzen zum Beutefang aus verschiedenen Seidentypen hergestellte Fangvorrichtungen. Die bekannteste dieser Strukturen ist sicherlich das Radnetz. Zum Bau dieser Netze verwenden Echte Radnetzspinnen (*Araneidea*) fünf verschiedene Seidenarten, wobei jeder Seidentyp eine wichtige und spezialisierte Aufgabe innerhalb des Netzes zukommt. Zählt man noch die für die Konstruktion des Eisacks sowie die für das Umwickeln der Beute verwendeten Seiden hinzu, sind weibliche Individuen der Echten Radnetzspinnen somit zur Produktion von sieben verschiedenen Seidenarten in der Lage<sup>[21,22]</sup>. Die meisten wissenschaftlichen Erkenntnisse hinsichtlich Aufbau, Eigenschaften und Funktion existieren für die zwei Hauptseidenarten der Radnetze. Die in der Großen Ampullendrüse gebildete Seide wird für die Konstruktion des Rahmes und der Speichen der Radnetze sowie als Abseilfaden, um sich vor Gefahren in Sicherheit zu bringen, verwendet. Diese Seide wird entweder aufgrund der sie bildenden Drüse als Große Ampullenseide (MAS, von engl. „*major ampullate silk*“) oder aufgrund ihrer Funktion auch als Abseilfaden bezeichnet. Ihre relativ unkomplizierte Zugänglichkeit (verglichen mit anderen Spinnenseidenarten) und ihre herausragenden mechanischen Eigenschaften<sup>[23]</sup> führten dazu, dass die Seide des Abseilfadens die bekannteste und bisher auch am gründlichsten erforschte Spinnenseide ist<sup>[24,25]</sup>. Eine zweite, weniger gut erforschte Hauptseidenart der Radnetze ist die Flagelliformseide, die die Fangspirale bildet und die sich durch eine außergewöhnliche Elastizität und Dehnbarkeit auszeichnet<sup>[14,26,27]</sup>.

### 1.1.2 Seidenähnliche Proteinstrukturen bei Muscheln

Neben den zahlreichen Vertretern der Arthropoden verwenden auch Lebewesen anderer Klassen extrakorporale Proteinstrukturen, die aufgrund ihrer Funktion und ihres Aufbaus mit Seiden vergleichbar sind und deshalb als seidenähnlich bezeichnet werden können. So besitzen einige Muschelarten die Fähigkeit, ein auf Proteinen basierendes, extrakorporales Gebilde (den sog. Muschelbyssus) herzustellen. Bei den zugrunde liegenden Proteinen handelt es sich zwar nicht um Seiden nach den in 1.1.1 genannten, „klassischen“ Definitionen, jedoch werden sie aufgrund ihrer natürlichen Funktion, ihres makroskopischen sowie molekularen Aufbaus und ihrer in der Antike erfolgten

Verwendung als Rohmaterial zur Herstellung exklusiver Textilien häufig ebenfalls als „Muschelseide“ bezeichnet<sup>[28,29]</sup>. Der Byssus selbst dient bestimmten, in Gezeitenzonen sesshaften Muschelarten als Verankerungs- und Schutzmechanismus. Mit Hilfe der zwei bis fünf Zentimeter langen Byssusfäden befestigen sich die Muscheln an Steinen, Pfählen oder auch anderen Muscheln in ihrer direkten Umgebung<sup>[28,30]</sup>. Über diese reine Verankerungsfunktion hinaus bietet der Byssus zusätzlich mechanischen Schutz, da er als Kompositmaterial mit sich graduell entlang der Fadenachse ändernden mechanischen Eigenschaften die durch Wellen und Meeresströmungen verursachten, auf die Muschel einwirkenden Kräfte größtenteils absorbieren kann<sup>[31]</sup>. Aufgrund der interessanten mechanischen Eigenschaften und seines Kompositmaterialcharakters rückt auch der Muschelbyssus zusehends in den Fokus von Natur- und Materialwissenschaften<sup>[31-43]</sup>.

### 1.2 Aufbau von Seiden und seidenähnlichen Strukturen

Sowohl Arthropodenseiden als auch der seidenähnliche Muschelbyssus sind komplexe, aus Proteinen und z. T. zusätzlichen Biopolymeren aufgebaute Kompositmaterialien. Ein ganzheitliches Verständnis der Eigenschaften dieser faszinierenden Materialien setzt sowohl die Erforschung des makroskopischen Aufbaus, eine umfassende Charakterisierung der beteiligten Einzelkomponenten und deren Interaktionen sowie in besonderem Maße auch die Aufklärung der Struktur-Funktions-Beziehungen der beteiligten Proteine voraus.

Trotz ihrer verschiedenen Ursprünge, ihrer konvergenten Evolution und ihrer breiten Verteilung über verschiedene Stämme des Tierreichs weisen Seiden und seidenähnliche Proteine sowohl auf makroskopischer als auch molekularer Ebene einige gemeinsame Prinzipien hinsichtlich Aufbau und Struktur auf. Dieser Abschnitt beschäftigt sich mit diesen gemeinsamen Prinzipien sowie den daraus bisher ableitbaren Erkenntnissen zu Struktur-Funktions-Beziehungen.

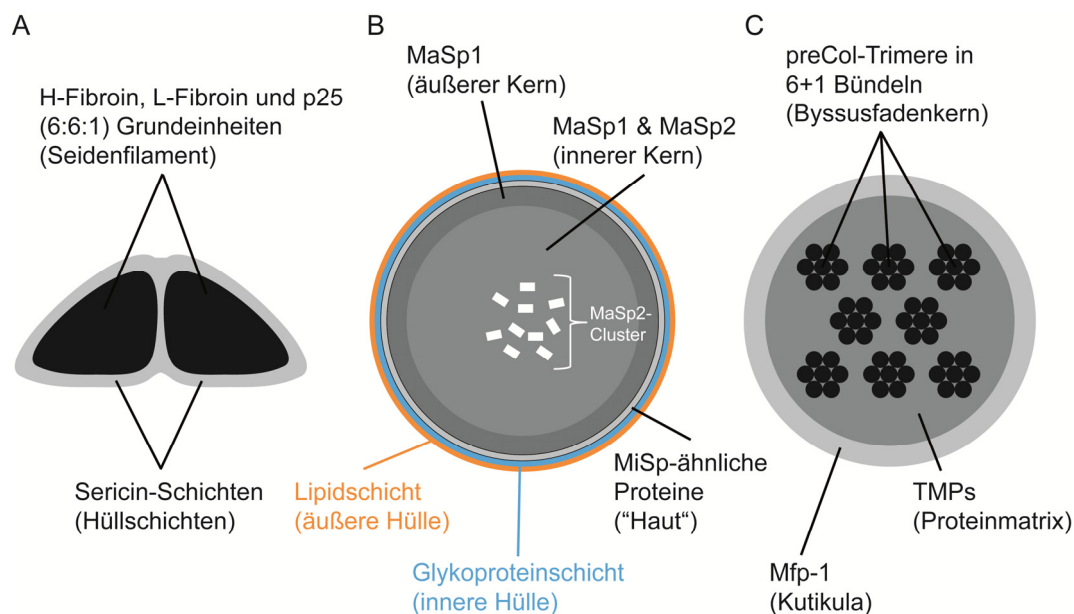
### 1.2.1 Makroskopischer Kern-Hülle-Aufbau

Sowohl die Seidenfäden der Insekten und der Webspinnen als auch der seidenähnliche Muschelbyssus sind komplexe, aus zahlreichen Komponenten bestehende Systeme. Das Zusammenspiel verschiedener Proteine führt zur Ausbildung hierarchischer Strukturebenen und äußert sich makroskopisch in der Ausbildung einer Kern-Hülle-Struktur. Der fibrilläre Kern, der für die mechanischen Eigenschaften des Gesamtfadens bestimmend ist, entsteht durch die Assemblierung der eigentlichen Seidenproteine bzw. im Fall des Muschelbyssus durch die Assemblierung der seidenähnlichen Moleküle (preCols, vgl. Kapitel 1.3.2). Dieser Kern ist mit weiteren Proteinen sowie teilweise auch anderen Biopolymeren (Lipide, Zuckerreste) beschichtet. Die dadurch entstehende Hülle übernimmt hauptsächlich Schutzfunktionen und verhindert eine Beschädigung des Kerns durch abiotische und biotische Einflüsse.

Der Kern des Seidenfadens von Larven des Maulbeerspinners *B. mori* besteht aus zwei gleichartigen Filamenten, die jeweils aus drei verschiedenen Proteinen aufgebaut sind: einer leichten und einer schweren Fibroinkette (L- und H-Fibroin, von engl. *light* und *heavy chain* fibroin) und dem Glykoprotein P25<sup>[3]</sup>. H- und L-Fibroin bilden ein disulfidverbrücktes Heterodimer<sup>[44]</sup> und 3 dieser Dimere mit P25 zusammen bilden die elementare Grundeinheit des Fadens (6:6:1 Stöchiometrie)<sup>[45]</sup>. Nach Assemblierung der Grundeinheiten werden jeweils zwei der entstandenen Filamente mittels eines zur Gruppe der Sericine gehörenden Proteins miteinander verklebt. Bei letzteren handelt es sich um repetitive, glykosylierte Proteine mit Molekulargewichten zwischen 24 kDa und 400 kDa, die einen Serinanteil von bis zu 40 % aufweisen<sup>[3,46]</sup>. Die Klebeeigenschaften der Sericine beruhen v. a. auf zwischen ihren Serinresten und den Fibroinfilamenten entstehenden Wasserstoffbrückenbindungen<sup>[3]</sup>. Das entstandene Doppelfilament, das den Kern der *B. mori* Kokonseide bildet, wird wiederum von mehreren, aus unterschiedlichen Sericinen gebildeten Hüllschichten ummantelt<sup>[3]</sup>. Diese schützen nicht nur den Fadenkern, sondern vermitteln ebenfalls Klebeeigenschaften, so dass die fertigen Kokonseiden analog zu den Fibroinfilamenten wiederum mit sich selbst sowie mit Umgebungsgegenständen verklebt werden können<sup>[3]</sup> (Abb. 1-2 A).

Das makroskopische Bauprinzip der Köcherfliegen-seide ist dem der Kokonseide der Schmetterlinge sehr ähnlich. Der Fadenkern besteht ebenfalls aus Filamenten, welche

aus H- und L-Fibroin aufgebaut sind<sup>[3,47,48]</sup>. Im Gegensatz zu den Schmetterlingsseiden weisen die Grundeinheiten der Kernfilamente jedoch kein P25 bzw. ein diesem homologes Protein auf<sup>[3,48]</sup>. Zwischen den einzelnen Filamenten sowie als Hülle um diese herum existieren wie bei den Schmetterlingsseiden Schichten aus Proteinen, die aufgrund ihrer Analogie zu den Sericinen der Schmetterlingsseiden ebenfalls als Sericin(e) bezeichnet werden<sup>[47,48]</sup> (Abb. 1-2 A).



**Abb. 1-2: Schematische Darstellung von Querschnitten verschiedener Seidenfäden und eines Muschelbyssusfadens. (A)** Fäden der Kokonseide von Vertretern der Schmetterlinge und Köcherfliegen bestehen aus miteinander verklebten Fibroinfilamenten (H- und L-Fibroin sowie je nach Ordnung weitere Proteine) und sind von mehreren Sericinschichten umhüllt. **(B)** Der in zwei Bereiche unterscheidbare Kern der Abseilfäden Echter Radnetzspinnen besteht aus homogen verteiltem MaSp 1 und inhomogen verteiltem MaSp 2. Die Hülle des Abseilfadens besteht aus 3 Schichten: einer als Haut bezeichneten Proteinschicht, die von MiSp-ähnlichen Proteinen (MiSp selbst wird zum Bau einer als Schablone für die Konstruktion der Fangspirale des Radnetzes dienenden Hilfspirale genutzt) gebildet wird sowie einer Glykoprotein- und einer abschließenden Lipidschicht. **(C)** Der Kern des Muschelbyssusfadens besteht aus 6+1 Bündeln von preCol-Trimern, die in eine aus Matrixproteinen (TMPs) bestehende Matrix eingebettet sind. Die Kutikula besteht hauptsächlich aus Mfp-1.

Die Seidenfäden von Vertreter der Hautflügler bestehen aus mehreren unterschiedlichen Proteinen. So können in Honigbienen, Hummeln, Hornissen und Ameisen vier in Zusammenhang mit der Seidenproduktion stehende Proteine mit Molekulargewichten von 29 kDa bis 60 kDa nachgewiesen werden<sup>[4]</sup>. Diese assemblieren zu tetrameren *coiled-coils*<sup>[49]</sup> und bilden den strukturellen Kern der Hautflüglerseiden. Der Fadenkern wird vermutlich ebenfalls von einer aus weiteren Proteinen bestehenden Hülle umgeben, da,



je nach Spezies, noch 1–2 weitere Proteine detektiert wurden, denen eine den Sericinen ähnliche Rolle zugeschrieben werden kann<sup>[4]</sup>.

Ähnlich den Insektenseiden sind auch einige Spinnenseiden komplexe Kompositmaterialien, die auf makroskopischer Ebene einen Kern-Hülle-Aufbau zeigen. Der strukturelle Kern des Abseilfadens Echter Radnetzspinnen kann in zwei Bereiche unterschieden werden und besteht aus Proteinen der beiden Spidroinklassen MaSp1 und MaSp2 (von engl. „major ampullate spidroin“) <sup>[9,50]</sup>. Während MaSp1 als Hauptkomponente des Fadenkerns homogen verteilt vorliegt und sich sowohl im inneren wie äußeren Kern findet, bildet MaSp2 Cluster, die auf das Zentrum des inneren Fadenkerns beschränkt sind<sup>[51]</sup>. Zudem zeigt das Verhältnis von MaSp1 zu MaSp2 speziesabhängige Unterschiede<sup>[52]</sup>. Der MAS-Fadenkern wird von einer Hülle umgeben, deren Aufbau komplexer als bei den Insektenseiden ist und die aus drei unterscheidbaren Schichten besteht. Der Fadenkern wird von einer als „Haut“ bezeichneten Schicht umschlossen, die aus einem Protein besteht, welches dem Kleinen Ampullenseidenprotein (MiSp, von engl. „minor ampullate spidroin“) ähnelt<sup>[51]</sup>. Seide aus letzterem wird von Radnetzspinnen normalerweise zum Bau einer (als Schablone dienenden) Hilfsspirale für die Konstruktion der eigentlichen Fangspirale des Radnetzes verwendet<sup>[20,25,53,54]</sup>. Nach dieser „Haut“ folgt noch eine aus Glykoproteinen bestehende Lage sowie eine abschließende Lipidschicht<sup>[51]</sup>. Die komplexe Hülle übernimmt mehrere Funktionen. Sie dient in erster Linie dem Schutz des Fadenkerns vor biotischen und abiotischen Umwelteinflüssen, übernimmt aber auch Aufgaben bei der sozialen Kommunikation zwischen Radnetzspinnen, da die Lipide der äußersten Hüllschicht auch als Träger von Pheromonen in Frage kommen<sup>[55]</sup> (Abb. 1-2 B).

Der bei den Seidenfäden der Arthropoden zu findende Kern-Hülle-Aufbau ist auch im Bauplan des seidenähnlichen Muschelbyssus erkennbar. Jeder Byssusfaden umfasst insgesamt bis zu 10 Strukturproteine und Enzyme<sup>[43]</sup>. Der strukturelle Kern des Byssusfadens, der makroskopisch für die mechanischen Eigenschaften der Fäden verantwortlich ist, besteht aus den drei kollagenähnlichen Proteinen preCol-D, preCol-P und preCol-NG, <sup>[31,56-58]</sup> (von engl. „pre-pesinized collagen“, vgl. Kapitel 1.3.2). Innerhalb des Fadenkerns formen die höchstwahrscheinlich als Homotrimere vorkommenden preCols 6+1 Bündel mit einer lokal begrenzten, flüssigkristallinen (smektischen) Ordnung. Die flächig angeordneten Bündel sind ihrerseits in eine Matrix, die die einzelnen preCol-Bündel

voneinander trennt, eingebettet<sup>[59,60]</sup>. Die Matrix selbst wird von Vertretern einer als TMP (von engl.: „*thread matrix proteins*“) bezeichneten Proteinfamilie gebildet<sup>[60]</sup>. Der strukturelle Kern wird von einer schützenden Hülle, der Kutikula, welche aus dem Protein Mfp-1 (von engl. „*mussel foot protein 1*“) sowie Lipiden und Wachsen besteht, umschlossen<sup>[37,41,61,62]</sup> (Abb. 1-2 C).

### 1.2.2 Molekularer Aufbau

Die den Arthropodenseiden zugrunde liegenden Proteine sowie die preCols des Muschelbyssus zeigen nicht nur im makroskopischen Bauplan Ähnlichkeiten. Auch auf molekularer Ebene sind Gemeinsamkeiten in der Primärstruktur der jeweiligen Proteine zu finden.

Der größte Teil der Seidensequenz (der deshalb oft auch als Kerndomäne bezeichnet wird) umfasst Abfolgen sich wiederholender und für das jeweilige Protein charakteristischer Konsensusmotive, die häufig so angeordnet sind, dass sich hydrophobe und hydrophile Bereiche in einer blockcopolymerähnlichen Weise abwechseln<sup>[31,63]</sup>. Die deutlich kürzeren, terminalen Domänen der Fibroine, Spidroine und preCols zeigen hingegen eine nichtrepetitive Aminosäuresequenz. Eine weitere Gemeinsamkeit ist, dass die Konsensusmotive der Kerndomänen aus einigen wenigen, meist nichtessentiellen Aminosäuren zusammengesetzt sind, was in einem überdurchschnittlich hohen prozentualen Anteil der entsprechenden Aminosäuren resultiert.

#### 1.2.2.1 Repetitive Kerndomänen

Die repetitive Kernsequenz der H-Fibroine von Schmetterlingen und Köcherfliegen weisen einen überproportional hohen Anteil der Aminosäuren Glycin und Serin, sowie bei Schmetterlingen zusätzlich noch Alanin, auf<sup>[5]</sup>. Der Kern des *B. mori* H-Fibroins wird von (GX)<sub>n</sub>- und GX<sup>1</sup>X<sup>2</sup>G-Sequenzmotiven (X ist vorwiegend eine der Aminosäuren A, S oder V) dominiert. Die beiden Arten von Sequenzmotiven sind so angeordnet, dass zwei Arten von Clustern unterschieden werden können: das (GX)<sub>n</sub>-Sequenzmotiv bildet sowohl GAGAGS-Abfolgen mit 2 bis 15 als auch GX<sup>1</sup>GX<sup>2</sup>GX<sup>3</sup>-Abfolgen (X ist meist eine der Aminosäuren A, V oder Y) mit 2 bis 5 Wiederholungen aus, und das GX<sup>1</sup>X<sup>2</sup>G-Sequenzmotiv

folgt meist auf das erste GAGAGS-Motiv innerhalb des Clusters. Diese hydrophoben Module können zu zwölf übergeordneten Modulen, die jeweils von 43 Aminosäuren umfassende, hydrophileren Sequenzen unterbrochen werden, zusammengefasst werden<sup>[3,63]</sup>. Wie bei *B. mori* findet man auch bei anderen Schmetterlingen diesen charakteristischen, modularen Aufbau der H-Fibroinsequenz, wobei Anzahl der Module sowie Art und evolutionäre Konservierung der darin vorkommenden Wiederholungssequenzen jedoch von Spezies zu Spezies variieren<sup>[3]</sup>.

Die repetitiven Kernsequenzen der H-Fibroine der Köcherfliegen unterscheiden sich von denen der Schmetterlinge hauptsächlich darin, dass alaninreiche Sequenzmotive  $((GX)_n$  mit  $X=A$  oder  $A_n$ ) fehlen (Tab. 1-1). Stattdessen sind neben den durch  $X^1X^2GPG$ -Motiven geprägten glycinreichen Bereichen v. a.  $(SX)_n$ -Sequenzabfolgen zu finden. Darüber hinaus ist die Anzahl unterscheidbarer Konsensusmotive größer als bei den H-Fibroinen der Schmetterlinge, die Sequenzen sind stärker konserviert und Muster einzelner Konsensusmotive werden deutlich seltener wiederholt, was letztendlich auch die Gruppierung dieser Motive zu übergeordneten Modulen erschwert. Unterschiede zwischen den Seidenproteinen einzelner Gattungen äußern sich bei Köcherfliegeneseiden deshalb eher in einer variablen Anordnung der Motive denn in Sequenzunterschieden innerhalb der Konsensusmotive<sup>[3]</sup>.

Der repetitive Kern der 4 paralogen, bei allen Vertretern der Hautflüglerüberfamilien *Vespoidea* und *Apoidea* vorkommenden Seidenproteine weist einen überdurchschnittlich hohen Anteil der Aminosäuren Alanin, Serin, Glutaminsäure (E), Leucin (L), Lysin (K), Glutamin (Q) und Threonin (T) auf. Charakteristisch ist für diese Proteine die sich wiederholende Abfolge eines Heptapeptidmotivs, wobei sich auf Position 1 und 4 hydrophobe Aminosäuren mit kleinen Seitenketten (meist A und S) befinden<sup>[4]</sup> (Tab. 1-1).

Die repetitiven Kernsequenzen der Proteine der beiden Spidroinklassen MaSp1 und MaSp2 der Echten Radnetzspinnen weisen analog den Fibroinen der Insektenseiden ebenfalls einen hohen Anteil der Aminosäuren Alanin, Glycin und Serin auf und sind in Form sich wiederholender Sequenzmotive angeordnet. Die Kernsequenz der MaSp1-Proteine beinhaltet Abfolgen alaninreicher  $((A)_{4-12} / (GA)_n$ ) und glycinreicher Motive (GGX, mit  $X = \text{Tyrosin (Y), L oder Q}$ ). Die MaSp2-Kernsequenzen beinhalten genauso  $(A)_{4-12}$ - und  $(GA)_n$ -Sequenzmotive, aber Glycin tritt typischerweise nicht mehr als GGX-,

sondern in Form von  $X^1X^2$ GPG-Motiven ( $X = Q, G$  oder  $Y$ ) auf<sup>[20,64,65]</sup> (Tab. 1-1). Dies verleiht MaSp2 einen etwas hydrophileren Charakter als MaSp1. Die Funktion der in den repetitiven Kernsequenzen der Insektenseidenproteine vorkommenden, hydrophilen Abstandhalter-Sequenzmotive (engl. *spacer*) wird bei den MAS durch unterschiedliche Mengenverhältnisse der beiden Spidroinklassen übernommen<sup>[51,63]</sup>.

Die Kernsequenzen der preCol-Moleküle des Muschelbyssus ähneln den Arthropodenseidenproteinen durch ihren überdurchschnittlich hohen Anteil der Aminosäuren Glycin und Alanin. Die repetitive Kernsequenz gliedert sich in zwei voneinander unterscheidbare Bereiche: einen glycinreichen Teil sowie zwei amino- und carboxyterminal zu diesem liegende Sequenzbereiche, die neben einem hohen Glycin- zusätzlich noch durch einen hohen Alaninanteil gekennzeichnet sind. Der glycinreiche Teil ist durch die strikte Abfolge eines einzigen Konsensusmotivs, dem für fibrilläre Kollagene charakteristischen Tripeptidmotiv  $GX^1X^2$  ( $X^1$  und  $X^2$  sind häufig Prolin (P) bzw. 4-trans-Hydroxyprolin), charakterisiert<sup>[31,56-58]</sup>, und wird deswegen auch bei den preCols allgemein als Kollagendomäne bezeichnet. Die amino- und carboxyterminal anschließenden Sequenzen, die aufgrund ihrer Position relativ zur kollagenartigen Kerndomäne auch als flankierende Regionen bezeichnet werden, unterscheiden sich deutlich zwischen den drei preCols und sind aus Konsensusmotiven zusammengesetzt, die charakteristisch für Fibroine, Spidroine und andere weit verbreitete Strukturproteine wie z.B. Elastin und Abductin sind<sup>[31,56-58]</sup> ( Tab. 1-1, und Kapitel. 1.3.2 für detailliertere Beschreibungen).

**Tab. 1-1: Übersicht der Konsensusmotive innerhalb der repetitiven Sequenzen ausgewählter Seidenproteine und Muschelbyssus preCols.** Konsensusmotive innerhalb der repetitiven Kernsequenzen von Fibroinen, Spidroinen und preCols. Abfolgen dieser Motive neigen zur Ausbildung verschiedener Proteinkonformationen, die wiederum für bestimmte Eigenschaften und Funktionen der jeweiligen Seide von Bedeutung sind. Hochgestellte Zahlen: Nummerierung der Positionen, an denen verschiedene Aminosäuren vorkommen können.

Konsensusmotiv:	Vorkommen:	Besonderheiten:	Konformation:	Rolle/Funktion:
<b>A<sub>n</sub></b>	MAS <sup>[20,64,65]</sup>	n = 4-12 <sup>[20,64,65]</sup>	β-Faltblätter in nanokristalliner Anordnung <sup>[20,25,66-72]</sup>	→ Zugfestigkeit & Steifigkeit
	preCol-NG ( <i>flanks</i> ) <sup>[31,58]</sup>	n = 3, auf aminoterminaler flank beschränkt <sup>[31,58]</sup>	β-Faltblätter, vermutlich parallel <sup>[73]</sup>	
	preCol-D ( <i>flanks</i> ) <sup>[31,57]</sup>	substituiert mit großen hydrophoben AS <sup>[31,57]</sup>	antiparallele β-Faltblätter in kristalliner Anordnung <sup>[73]</sup>	→ preCol-D: Rolle während des Selbstheilungsprozesses im distalen Teil von Byssusfäden
<b>(XG)<sub>n</sub></b>	H-Fibroine (Schmetterlinge) <sup>[63]</sup>	X = A, Y, V <sup>[63]</sup>	β-Faltblätter in kristalliner Anordnung <sup>[3,20,31,57,58,70-75]</sup>	→ Zugfestigkeit & Steifigkeit
	MAS <sup>[20,64,65]</sup>	X = A <sup>[20,64,65]</sup>		→ MAS: Zwischenbereich zwischen β-Nanokristallen und nicht-kristallinen Bereichen
	preCol-D ( <i>flanks</i> ) <sup>[31,57]</sup>	X = A / n = 2 <sup>[31,57]</sup>		
	preCol-NG ( <i>flanks</i> ) <sup>[31,58]</sup>	meist: GAGA, GAGS <sup>[31,58]</sup>		
<b>(GX<sup>1</sup>GX<sup>2</sup>GX<sup>3</sup>)<sub>n</sub></b>	H-Fibroine (Schmetterlinge) <sup>[3,63]</sup>	häufig: X <sup>1</sup> , X <sup>2</sup> = A, X <sup>3</sup> = S <sup>[3,63]</sup>	β-Faltblätter in kristalliner Anordnung <sup>[3,76]</sup>	→ Zugfestigkeit & Steifigkeit
<b>(SX)<sub>n</sub></b>	H-Fibroine (Köcherfliegen) <sup>[3]</sup>	X ≠ A (keine A-reichen Motive vorhanden) <sup>[3]</sup>	β-Faltblätter in kristalliner Anordnung <sup>[3]</sup>	
<b>(XG<sub>m</sub>)<sub>n</sub></b>	H-Fibroine (Köcherfliegen) <sup>[3]</sup>	oft: X = L, m = 2 <sup>[3]</sup>	Typ-II β-Schleifen oder helikale Konformationen (v. a. 3 <sub>1</sub> -helikal) <sup>[3,20,31,56-58]</sup>	→ Elastizität & Dehnbarkeit
	MaSp1 <sup>[20,64,65]</sup>	X = A, L, Q / m = 2 <sup>[20,64,65]</sup>		
	preCol-NG ( <i>flanks</i> ) <sup>[31,58]</sup>	häufig: X = L, A / m = 1-3 <sup>[31,58]</sup>		→ preCol-NG: Rolle während des Selbstheilungsprozesses im distalen Teil von Byssusfäden
	preCol-D ( <i>flanks</i> ) <sup>[31,57]</sup>	m = 2 <sup>[31,57]</sup>		
	preCol-P ( <i>flanks</i> ) <sup>[31,56]</sup>	häufig: X = V, I / m = 2 <sup>[31,56]</sup>		
<b>(X<sup>1</sup>X<sup>2</sup>X<sup>3</sup>PG)<sub>n</sub></b>	H-Fibroine (einige Vertreter der Köcherfliegen) <sup>[3]</sup>	meist: X <sup>3</sup> = G <sup>[3]</sup>	Typ-II & Typ-I β-Schleifen sowie <i>random coil</i> Bereiche <sup>[77,78]</sup> Abfolge vieler β-Schleifen kann zur Bildung einer sog. „β-Spirale“ führen <sup>[26,79]</sup>	→ Elastizität
	MaSp2 <sup>[20,64,65]</sup>	X <sup>1</sup> & X <sup>2</sup> = Q, G, Y, X <sup>3</sup> = G <sup>[20,64,65]</sup>		→ Entstehung einer elastischen (entropisch bedingten) Rückstellkraft bei Fadendehnung
	Flagelliform-Seide <sup>[14,20,26,27]</sup>	X <sup>1</sup> = G, X <sup>3</sup> = G <sup>[14,26,27]</sup>		
	preCol-P ( <i>flanks</i> ) <sup>[31,56]</sup>	X = noch mind. 2 weitere G <sup>[31,56]</sup>		
<b>GX<sup>1</sup>X<sup>2</sup></b>	preCols (Kollagendomäne) <sup>[31,56-58]</sup>	oft: X <sup>1</sup> = P, X <sup>2</sup> = Hyp <sup>[31,56-58]</sup>	kollagenartige Tripelhelix <sup>[31,56-58]</sup>	→ strukturelle & mechanische Integrität der Byssusfäden → Assemblierung der preCols zu Tripelhelices
<b>(A/SX<sup>1</sup>X<sup>2</sup>A/SX<sup>3</sup>X<sup>4</sup>X<sup>5</sup>)<sub>n</sub></b>	Seidenproteine (Vertreter der <i>Apoidea</i> u. <i>Vespoidea</i> ) <sup>[4]</sup>	A bzw. S an Pos. 1+4 <sup>[4]</sup>	tetramere <i>coiled coils</i> <sup>[4]</sup>	

Ähnliche Abfolgen und Wiederholungen verwandter Konsensusmotive tendieren im fertigen Faden auch zur Ausbildung ähnlicher Sekundärstrukturen.

Die alaninreichen Bereiche der H-Fibroine der Schmetterlinge und der MAS der Echten Radnetzspinnen sowie die serinreichen  $(SX)_n$ -Motive der Köcherfliegen-seidenproteine bilden bevorzugt  $\beta$ -Faltblattstrukturen aus<sup>[3,66-71]</sup>. Gleiches gilt für Sequenzmotive, in denen Glycin hauptsächlich mit Alanin oder Serin alterniert  $((XG)_n$  oder  $GX^1X^2G$ )<sup>[70,71,74,75]</sup>. Die  $\beta$ -Faltblätter können sowohl parallel wie auch senkrecht zur Fadenachse angeordnet sein (vgl. Abb. 1-1 B) und erlauben eine dichte Packung der Moleküle<sup>[5]</sup>, was bei  $\sim 40\%$  der entsprechenden Sequenzbereiche zur Bildung kristalliner Regionen führt<sup>[20,71]</sup>. Lediglich die Seidenproteine von Vertretern der Hautflüglerüberfamilien *Apoidea* und *Vespoidea* weichen von diesem Muster ab, da in den alaninreichen Sequenzbereichen vorrangig eine *coiled-coil* Konformation anstelle einer  $\beta$ -Faltblattkonformation nachgewiesen werden kann<sup>[5,49]</sup>, die mit dem charakteristischen Verteilungsmuster der Alanine innerhalb der Heptapeptidwiederholungen in Verbindung steht.

Die glycinreichen Aminosäuresequenzen der Seidenproteine bilden hingegen Sekundärstrukturmotive aus, die mit elastischen Eigenschaften in Verbindung gebracht werden. Sowohl  $(XG_m)_n$ - wie auch  $(X^1X^2GPG)_n$ -Motive bilden häufig  $\beta$ -Schleifen (hauptsächlich Typ-II, aber auch Typ-I), die, wenn sie direkt hintereinander auftreten, eine sog.  $\beta$ -Spirale bilden können<sup>[77-79]</sup>. Eine bestimmte Variante der  $(XG_m)_n$ -Motive,  $(XGG)_n$ , ist zudem an der Ausbildung  $3_1$ -helikaler Konformationen beteiligt<sup>[3,20,31,56-58]</sup>. Im Gegensatz zu den geordneten und dicht gepackten  $\beta$ -Faltblattkonformationen weisen Molekülbereiche mit Schleifen- sowie helikalen Konformationen meist eine lockerere Packung und einen geringeren Grad an Orientierung sowie Ordnung auf<sup>[3,14,20,27,80]</sup>. Einen Sonderfall stellen die kollagenähnlichen  $(GX^1X^2)_n$ -Motive der preCols dar, die eine Polyprolin-II-helixähnliche (PPII-Helix) Konformation einnehmen und wie die fibrillären Vertebratenkollagene auch zu dicht gepackten, rechtsgängigen Tripelhelices assemblieren<sup>[36,81]</sup>.

### 1.2.2.2 Nichtrepetitive Amino- und Carboxytermini

Die großen, repetitiven Kerndomänen der hier beschriebenen (sowie die der meisten anderen) Arthropodenseidenproteine und der preCols werden amino- und carboxyterminal von nichtrepetitiven Domänen abgeschlossen. Diese sind in der Regel hydrophiler

als die jeweiligen repetitiven Kernsequenzen und beinhalten Aminosäuren, über deren Seitenketten sowohl kovalente als auch nichtkovalente intra- und intermolekulare Interaktionen möglich sind.

Die Termini der H-Fibroine der Schmetterlings- und Köcherfliegenseiden sind zwischen beiden Ordnungen konserviert. Die Aminotermini bestehen aus einer nichtrepetitiven, amphiphilen Aminosäuresequenz<sup>[3]</sup>. Der Aminoterminus des *B. mori* H-Fibroins umfasst 151 Aminosäuren und kann pH-Wert-abhängig sowohl in einer *random coil*- als auch einer antiparallelen  $\beta$ -Faltblattkonformation vorliegen<sup>[82,83]</sup>. Aufgrund der erwähnten Konservierung der aminoterminalen Sequenzen wird davon ausgegangen, dass diese sich auch in den H-Fibroinen anderer Schmetterlings- und Köcherfliegenseidenproteine konformationell ähnlich verhalten<sup>[82]</sup>. Die carboxyterminalen Domänen der H-Fibroine von Schmetterlingen und auch Köcherfliegen zeichnen sich v. a. durch das Auftreten von Cysteinresten an konservierten Positionen aus<sup>[3]</sup>. Der aus 50 Aminosäuren bestehende, nichtrepetitive Carboxyterminus des *B. mori* H-Fibroins beinhaltet drei solcher Cysteinreste<sup>[3]</sup>. Zwei davon bilden eine intramolekulare Disulfidbrücke, während der dritte eine Disulfidbrücke zu einem Cysteinrest des L-Fibroins ausbildet<sup>[3,44]</sup>. Die korrekte Bildung spezifischer Disulfidbrücken spielt bei der Biosynthese der Insektenseidenfäden eine entscheidende Rolle, da nur das kovalent verbrückte Heterodimer aus H- und L-Fibroin korrekt sekretiert werden kann<sup>[84]</sup>.

Innerhalb der beiden Hautflüglerüberfamilien *Apoidea* und *Vespoidea* sind die terminalen Domänen der *Vespoidea*-Seiden in der Regel länger als die der *Apoidea*-Seiden. Zwischen den beiden Überfamilien sind die Termini anders als zwischen verschiedenen Insekten- bzw. Spinnenseidenproteinen jedoch im Laufe der Evolution über den Grad der Sequenzwiedererkennung hinaus divergiert<sup>[4]</sup>. Generell weisen die Termini der Hautflüglerseidenproteine einen erhöhten Serinanteil auf, was diese, wie die nichtrepetitiven Domänen anderer Seidenproteine auch, polarer als die repetitive Kernsequenz macht. Innerhalb natürlicher Hautflüglerseiden lassen sich sowohl  $\alpha$ -Helix- wie auch  $\beta$ -Faltblattanteile nachweisen und es gibt Hinweise darauf, dass die serinreichen Termini der Proteine – ähnlich den repetitiven (SX)<sub>n</sub>-Motiven der H-Fibroine der Köcherfliegen (vgl. Kapitel 1.2.2.1) hauptsächlich für die Ausbildung der  $\beta$ -Faltblätter verantwortlich sind<sup>[85]</sup>.

Bei den MaSp der Webspinnen wird die repetitive Kernsequenz der Proteine durch ~ 100–150 Aminosäuren umfassende Termini umschlossen<sup>[6,65,86]</sup>. Ihre Sequenzen sind zwischen den verschiedenen Seidenarten und selbst zwischen phylogenetisch nur entfernt verwandten Spinnenarten stark konserviert<sup>[6,80,86-90]</sup>. Röntgenstrukturanalysen zeigen, dass der globuläre MaSp1 Aminoterminus der Spinne *Euprosthops australis* fünf  $\alpha$ -Helices umfasst, die in Form eines Helixbündels so angeordnet sind, dass sich die elf geladenen Aminosäuren innerhalb der Tertiärstruktur gruppieren und so einen negativ und einen positiv geladenen Molekülpol ausbilden können<sup>[91,92]</sup>. Die Tertiärstruktur des Carboxyterminus besteht ebenfalls aus fünf  $\alpha$ -Helices. Eine dieser Helices beherbergt einen (innerhalb der MaSp1 Proteine) stark konservierten Cysteinrest. Sowohl die einzigartige Tertiärstruktur als auch definierte Aminosäuren an konservierten Positionen spielen bei den Struktur-Funktions-Beziehungen der carboxyterminalen Domänen eine essentielle Rolle (vgl. Kapitel 1.2.3).

Die terminalen Domänen der preCols des Muschelbyssus zeichnen sich durch das Vorkommen von Histidinclustern sowie 3,4-Dihydroxyphenylalaninen (DOPA), welche post-translational aus Tyrosinen gebildet werden, aus<sup>[31,93]</sup>. Sowohl die Histidincluster als auch die DOPA-Reste dienen der intra- sowie intermolekularen Quervernetzung der preCols bzw. preCol-Bündel<sup>[93-100]</sup>.

Durch Variation von Art, Abfolge und Anzahl von Konsensusmotiven können sich innerhalb der Fibroine, Spidroine und preCols verschiedenste Sekundär- und Tertiärstrukturen (sowie ausgehend davon auch Quartärstrukturen) ausbilden. Diese durch die Aminosäuresequenz beeinflusste strukturelle Diversität wirkt sich auf die Ordnung, Ausrichtung und Packung der Moleküle bzw. einzelner Molekülbereiche aus und bestimmt so die makroskopischen (mechanischen) Eigenschaften des jeweiligen Seidenfadens bzw. des Muschelbyssus.

### 1.2.3 Struktur-Funktions-Beziehungen bei Seiden und seidenähnlichen Proteinen

Die makroskopischen Eigenschaften einzelner Arthropodenseiden und des Muschelbyssus stehen in enger Beziehung zur Primärstruktur der zugrunde liegenden Proteine, da diese Sekundär-, Tertiär- und Quartärstruktur und somit indirekt auch Ordnung,



Ausrichtung und Packung der Moleküle vorgibt, und durch ihren meist amphiphilen Charakter und das Vorhandensein bestimmter Aminosäuren weitere Eigenschaften wie die Löslichkeit der Proteine in den Drüsen, deren Unlöslichkeit im fertigen Seidenfaden oder auch die Haftfähigkeit der Seidenfäden selbst mitbestimmt.

Die repetitiven Kerndomänen der Arthropodenseidenproteine und der preCols enthalten einen hohen Anteil von Aminosäuren mit relativ kleinen Seitenketten (A, G, S) von intermediärer Hydrophobizität<sup>[4,5,20,31,56-58,64,65]</sup>. Aufgrund dieser Eigenschaften tragen die repetitiven Sequenzmotive zur Wasserlöslichkeit der Proteine vor und ihrer Unlöslichkeit nach der Assemblierung bei<sup>[5]</sup>. Darüber hinaus können die repetitiven Sequenzen in eine Vielzahl von Sekundärstrukturmotiven falten (Abb. 1-1 B & Kapitel 1.2.2). Eine periodische Abfolge dieser Sekundärstrukturen beeinflusst Ordnung und Packung der Seidenproteine bzw. preCols während der Assemblierung und trägt dadurch entscheidend zu den mechanischen Eigenschaften der Seiden bzw. des Byssus bei<sup>[5]</sup>. In ihrer löslichen Form weisen sowohl die repetitiven Sequenzmotive der H-Fibroine als auch die der MaSp vorwiegend helikale (amorphe) Konformationen auf. Bei Fibroinen liegt meist eine durch  $\beta$ -Schleifen (Typ-II) dominierte Konformation (auch als *Silk I*-Konformation bezeichnet) vor, während die repetitiven Kerndomänen der MaSp eine Mischung aus *random coil*, PPII- sowie  $\alpha$ -helikalen Konformationen zeigen<sup>[101-103]</sup>. Die kollagenartige Kernsequenz der preCols favorisiert eine PPII-helixähnliche Konformation, die bereits früh in den sekretorischen *Granulae* zur Ausbildung tripelhelikaler Homotrimere führt<sup>[36,43,59,104]</sup>. Während diese früh ausgebildete Quartärstruktur im fertig gebildeten Byssusfaden erhalten bleibt, führt der Spinnprozess bei Arthropodenseidenproteinen zu Konformationsänderungen. Sowohl bei den H-Fibroinen als auch den MaSp kommt es innerhalb der alaninreichen Sequenzen (bzw. den anstelle dieser bei Köcherfliegen- sowie einigen Schmetterlings-H-Fibroinen auftretenden serinreichen Sequenzen) zur Ausbildung von  $\beta$ -Faltblättern<sup>[3,63,66-71]</sup>. Die entstehenden  $\beta$ -Faltblätter ermöglichen eine sehr dichte und präzise Packung und Ausrichtung der Moleküle in den entsprechenden Sequenzbereichen. Dadurch können sich während des Verspinnens kristalline Bereiche (sog.  $\beta$ -Kristallite) mit Dimensionen von 2 x 5 x 7 nm, die hauptsächlich entlang der Fadenlängsachse ausgerichtet sind, innerhalb der Spinnenseidenfäden bilden<sup>[20,25,72]</sup>. Die alaninreichen Sequenzbereiche der Hautflüglerseiden stellen innerhalb der Fibroinsequenzen

eine Ausnahme dar, da sie aufgrund der charakteristischen Aminosäureabfolge der Heptapeptidwiederholungen sowie der dort zu findenden Verteilung der Alanine *coiled coils* (und somit eine  $\alpha$ -helikale Sekundärstruktur) anstelle von  $\beta$ -Faltblättern ausbilden<sup>[4,49]</sup>. Sequenzmotive mit hohem Glycinanteil (wie XGG oder  $X^1X^2$ GPG-Wiederholungen) neigen hingegen dazu, lockerer gepackte, nichtkristalline Konformationen einzunehmen. So bilden  $(XGG)_n$  Motive  $3_1$ -Helices und  $(X^1X^2GPG)$  bildet vorwiegend  $\beta$ -Schleifen, die, wenn sie in direkter Abfolge hintereinander liegen, eine sog.  $\beta$ -Spirale formen<sup>[3,20,77,79,80]</sup>. Motive, in denen sich Glycin mit Alanin oder Serin abwechselt  $((AG)_n$  oder  $(SG)_n$ ) tendieren jedoch wie alaninreiche Sequenzabschnitte eher zur Ausbildung von  $\beta$ -Faltblatt Strukturen und können wie letztere ebenfalls kristalline Bereiche bilden. Im Gegensatz zu den von den alaninreichen Motiven gebildeten  $\beta$ -Kristalliten sind diese Bereiche jedoch weniger dicht gepackt und besitzen eine größere räumliche Ausdehnung. Entsprechende, im Spinnenabseilfaden neben den  $\beta$ -Kristalliten vorhandene und nachweislich durch die  $(AG)_n$ -Motive gebildete kristalline Regionen weisen Dimensionen von 70–500 nm auf<sup>[20,25]</sup>. Die durch das H-Fibroinsequenzmotiv  $(GAGAGS)_n$  gebildeten  $\beta$ -Faltblätter können sich ebenfalls zu kristallinen Bereichen zusammenlagern. Ihre Dimensionen sind mit  $\sim 10$  nm lateraler und  $\sim 70$  nm axialen Ausdehnung<sup>[76]</sup> jedoch etwas kleiner als die der durch die  $(AG)_n$ -Sequenzmotive der MaSp gebildeten kristallinen Bereiche.

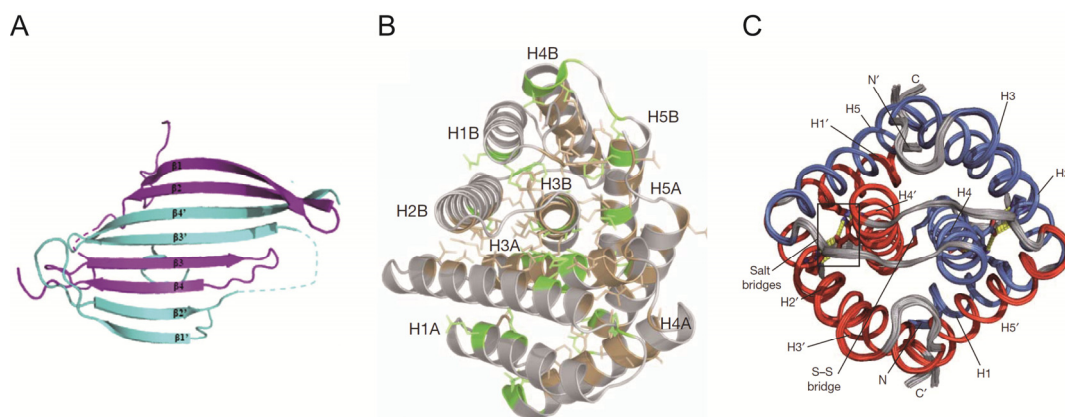
Die Kollagendomäne  $(GX^1X^2)$  der preCols ermöglicht die Ausbildung einer tripelhelikalen Quartärstruktur der preCol-Homotrimere und bestimmt somit in erster Linie auch deren mechanische Eigenschaften<sup>[105-110]</sup>. Die flankierenden Regionen der Kerndomäne unterscheiden sich signifikant zwischen den drei verschiedenen preCol-Molekülen und umfassen Konsensusmotive, die auch in einer Vielzahl anderer Strukturproteine vorkommen (Tab. 1-1 & Kapitel 1.3.2). Biochemische, biophysikalische sowie mechanische Untersuchungen lassen darauf schließen, dass die flankierenden Regionen hauptsächlich für die mechanisch unterschiedlichen Eigenschaften des distalen und proximalen Teils von Byssusfäden verantwortlich sind und auch bei der beobachtbaren Selbstheilung des distalen Abschnitts der Byssusfäden eine wichtige Rolle übernehmen könnten<sup>[59,73,94,106,108,110,111]</sup>.

Die Aufgaben der nichtrepetitiven, terminalen Domänen von Arthropodenseidenproteinen und preCols bestehen häufig darin, die Löslichkeit der Proteine vor, als auch deren korrekte Assemblierung während und nach der Fadenentstehung zu gewährleisten.

Die Sequenzen der nichtrepetitiven Amino- und Carboxytermini der H-Fibroine sind innerhalb der Schmetterlinge und Köcherfliegen konserviert. Eine rekombinant hergestellte Variante der aminoterminalen Domäne des *B. mori* H-Fibroins faltet *in vitro* bei leicht saurem pH-Wert (der auch während der Produktion der Kokonseide *in vivo* vorliegt) in ein aus vier Strängen bestehendes antiparalleles  $\beta$ -Faltblatt. Durch Domänen austausch (engl. *domain swapping*) der  $\beta$ -Stränge zweier H-Fibroinmoleküle bildet sich ein Homodimer (Abb. 1-3 A), welches dann mit einem weiteren Homodimer zu einem Homotetramer assembliert<sup>[82]</sup>. Durch diese Quartärstruktur kommen auf beiden Seiten des  $\beta$ -Faltblatts saure Aminosäurereste zu liegen. Diese dienen als pH-abhängiger Schalter, da ein weiteres Absinken des pH-Werts zu ihrer Protonierung und dadurch zum Verschwinden elektrostatischer Abstoßung führt. So sind eine dichte Packung sowie ausgeprägtere nichtkovalente Interaktionen zwischen den Tetrameren möglich<sup>[82]</sup>. Der nichtrepetitive Aminoterminus des *B. mori* H-Fibroins steuert somit die Oligomerisierung des Proteins und es darf aufgrund der Konservierung der H-Fibroine der Schmetterlinge und Köcherfliegen davon ausgegangen werden, dass dies eine generelle Funktion des Aminoterminus aller H-Fibroine ist. Darüber hinaus bindet der amphiphile Aminoterminus der Schmetterlings-H-Fibroine das Protein P25 über hydrophobe, nichtkovalente Wechselwirkungen. Die Interaktion mit dem hydrophileren P25 erleichtert den Transport und die Sekretion der schwer löslichen Fibroin-Heterodimere, die durch Disulfidverbrückung eines konservierten Cysteinrests im Carboxyterminus des H-Fibroins mit einem Cysteinrest des L-Fibroins entstehen<sup>[3,44,112]</sup>. Über die Termini vermittelte Quartärstrukturbildungen spielen somit bei Sekretions- wie auch Assemblierungsprozessen der Insektenseidenproteine eine wichtige Rolle.

Die aminoterminalen Domänen (NRN, von engl. „*non-repetitive N-terminal domain*“) des *E. australis* und des *Latrodectus hesperus* MaSp1 falten in eine bisher einzigartige Tertiärstruktur, ein aus fünf  $\alpha$ -Helices bestehendes Helixbündel<sup>[91,113]</sup> (Abb. 1-3 B). Diese Helixbündelstruktur ermöglicht eine räumliche Gruppierung der in der Sequenz vorhandenen geladenen Aminosäuren, so dass ein Ladungsdipol mit positiv und negativ geladenem Ende entsteht<sup>[91]</sup>. Bei neutralen pH-Werten und in Gegenwart von NaCl,

einer chemischen Umgebung, wie sie innerhalb der Großen Ampullendrüsen, in denen die MaSp gespeichert werden, vorliegt, ist NRN *in vitro* ein Monomer mit geringer Aggregationsneigung<sup>[113,114]</sup>. Eine Erniedrigung des pH-Werts (entspricht der Situation im Spinnkanal während des Spinnprozesses)<sup>[115]</sup> führt dann dazu, dass die Helixbündelstruktur intermolekulare, nichtkovalente Wechselwirkungen eingeht. Dies führt zu einer Dimerisierung von MaSp1, und aufgrund des Ladungsdipols der Tertiärstruktur erfolgt die Dimerisierung in einer antiparallelen Ausrichtung der beteiligten Monomere<sup>[91]</sup>. Die starke Konservierung zahlreicher Aminosäuren innerhalb der aminoterminalen Domänen der MaSp verschiedener Spinnenarten lässt darauf schließen, dass die Aufgabe letzterer generell darin liegt, einer frühzeitigen Aggregation der MaSp in der Drüse entgegenzuwirken und anschließend die antiparallele Multimerisierung der Proteine über nichtkovalente Wechselwirkungen während des Spinnprozesses zu steuern<sup>[91,116,117]</sup>.



**Abb. 1-3: Tertiär- und Quartärstrukturen terminaler Domänen von H-Fibroin und MaSp. (A)** Homodimer einer rekombinant hergestellten Variante des H-Fibroin Aminoterminus von *B. mori*, das durch Domänen-austausch einzelner  $\beta$ -Stränge gebildet wird. Türkis: Monomer A, violett: Monomer B. Abgedruckt aus dem Journal of Molecular Biology, Vol. 418, He Y. X. *et al.*, N-Terminal Domain of Bombyx mori Fibroin Mediates the Assembly of Silk in Response to pH Decrease, 197–207, Copyright (2012) **(B)** Antiparalleles Homodimer einer rekombinant hergestellten Variante des MaSp1-Aminoterminus, das durch nichtkovalente Interaktion der aus jeweils fünf  $\alpha$ -Helices bestehenden Helixbündel der Monomere gebildet wird. Konservierte Aminosäurereste, die in allen bekannten MaSp-Aminotermini vorkommen sind grün und diejenigen, die in > 60 % der MaSp-Aminotermini vorkommen, sind braun markiert. Abgedruckt mit Genehmigung der Macmillan Publishers Ltd: Nature (Askarieh, G. *et al.*, Nature 465 (7295), 236-238), Copyright (2012) **(C)** Paralleles Homodimer einer rekombinant hergestellten Variante des ADF3-Carboxyterminus (ADF3 ist ein MaSp2-Protein) von *A. diadematus*, das durch Domänen-austausch je einer  $\alpha$ -Helix der aus insgesamt fünf  $\alpha$ -Helices bestehenden Helixbündel sowie durch Ausbildung einer intermolekularen Disulfidbrücke entsteht. Blau: Monomer A, rot: Monomer B. Abgedruckt mit Genehmigung der Macmillan Publishers Ltd: Nature (Hagn, F. *et al.*, Nature 465 (7295), 239-243), Copyright (2012)

NMR-Untersuchungen des nichtrepetitiven Carboxyterminus (NRC, von engl. „*non-repetitive C-terminal domain*“) des MaSp2-Analogs ADF-3 von *A. diadematus* zeigen, dass auch diese Domäne in eine einzigartige, aus fünf  $\alpha$ -Helices bestehende Helixbündelkonformation faltet. Zwei NRC-Domänen interagieren durch Domänen austausch je einer der  $\alpha$ -Helices und bilden ein paralleles Dimer. Dieses wird durch Bildung einer stark konservierten, intermolekularen Disulfidbrücke zusätzlich stabilisiert<sup>[116,118]</sup> (Abb. 1-3 C). Unter den in der Spinndrüse vorherrschenden Bedingungen bilden die carboxyterminal verknüpften MaSp-Dimere mizellenartige, lösliche Strukturen<sup>[10,119,120]</sup>. Erst durch die während des Spinnprozesses auftretenden Änderungen der Ionenstärke und des pH-Werts sowie durch hinzukommende Scherkräfte ändert sich die Tertiärstruktur innerhalb der NRC. Es kommt zu einer Exposition hydrophober Bereiche, die die enge Packung, Ausrichtung und Assemblierung der MaSp einschließlich der Sekundärstrukturelemente ihrer repetitiven Domänen und damit die Bildung der kristallinen Bereiche im Seidenfaden ermöglicht<sup>[118]</sup>. Genau wie NRN ist auch NRC ein „molekularer Schalter“, der die Löslichkeit der MaSp innerhalb der Drüse fördert und wichtige Funktionen bei der Steuerung der Assemblierung übernimmt<sup>[116]</sup>.

Verglichen mit den nichtrepetitiven Termini der Arthropodenseidenproteine ist über Sekundär-, Tertiär- und Quartärstruktur der preCol-Termini und somit auch über ihre Struktur-Funktions-Beziehungen recht wenig bekannt. Generell sind die terminalen Domänen der drei preCols durch das Vorkommen histidinreicher Motive sowie konservierter Tyrosine, die zum Teil durch posttranslationale Modifikation zu DOPA hydroxyliert werden, gekennzeichnet<sup>[57,93,94]</sup>. Die Histidinreste können zusammen mit zweiwertigen Metallkationen koordinative Metall-Ligand-Komplexe, die zu einer stabilen, nichtkovalenten Quervernetzung der preCols führen, bilden<sup>[93-98]</sup>. Eine weitere Stabilisierung der Struktur findet durch die in den terminalen Domänen vorkommenden DOPA-Reste statt, da diese über eine Aryl-Aryl-Kupplung vermutlich intermolekulare di-DOPA Bindungen bilden<sup>[99,100]</sup>. Zusätzlich kann DOPA mit Metallkationen (allen voran  $\text{Fe}^{3+}$ ) ebenfalls Metall-Ligand-Komplexe bilden. Das Auftreten dieser Art der Quervernetzung scheint jedoch hauptsächlich auf die Kutikula (und damit nicht auf preCol-Moleküle) der Byssusfäden (Abb. 1-2 C & Kapitel 1.2.1) beschränkt zu sein<sup>[121]</sup>. Die preCol-Termini übernehmen somit vor allem Funktionen bei der Quervernetzung und Stabilisierung der bereits

während der Sekretion der preCols festgelegten flüssigkristallinen Orientierung der Moleküle.

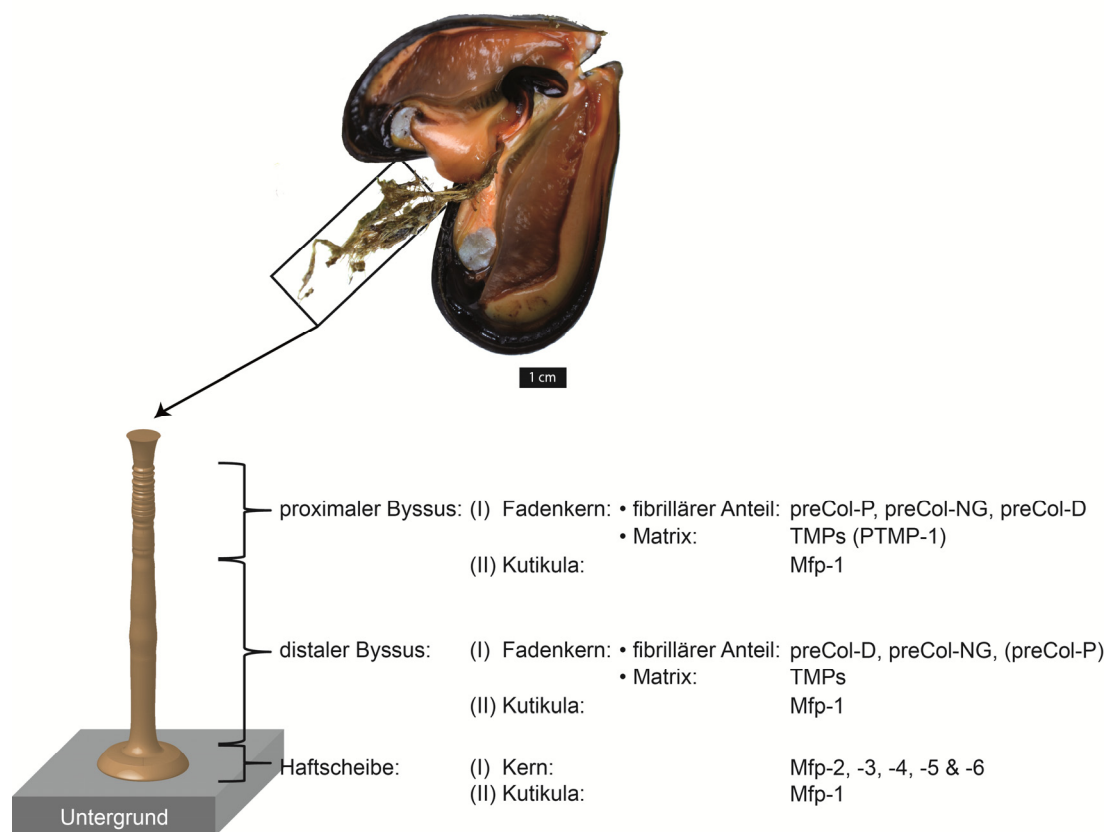
Trotz vieler Ähnlichkeiten auf makroskopischer und molekularer Ebene unterscheiden sich verschiedene Seidenarten sowie Seiden und Byssusfäden in ihren Eigenschaften. Um präziser verstehen zu können, worauf die spezifischen Eigenschaften sowie deren Unterschiede basieren, ist es unabdingbar, genaue Informationen über die Beziehungen zwischen den einzelnen Modulen (bzw. den darin enthaltenen Konsensusmotiven) und den daraus resultierenden und durch den Herstellungsprozess mit beeinflussten Konformationen zu erlangen. Nur eine umfassende Aufklärung dieser Struktur-Funktions-Beziehungen ermöglicht es, den Ursprung der herausragenden mechanischen und funktionalen Eigenschaften der hier beschriebenen extrakorporalen Strukturproteine erklären und das gewonnene Wissen als Grundlage zur Entwicklung neuartiger Biomaterialien nutzen zu können. Im Folgenden wird nun detaillierter auf den Aufbau des Muschelbyssus und die Struktur-Funktions-Beziehungen der zugrunde liegenden preCols, die Bestandteil dieser Arbeit waren, eingegangen.

### 1.3 Der Muschelbyssus

Muscheln der Familie *Mytilidae* sind sesshafte Mollusken, die die Gezeitenzonen von Küsten klimatisch gemäßigter Regionen bewohnen. Um ein durch Wellengang, Meeresströmungen und Gezeiten bedingtes Verdriften zu verhindern, nutzen Muscheln ein als Byssus bezeichnetes Verankerungssystem. Der Byssus kann morphologisch grob in drei Bereiche unterteilt werden: (i) in eine in die Byssusdrüse eingebettete Wurzel, (ii) in einen Stamm, der die nahtlose Fortsetzung der Wurzel darstellt und (iii) in die Byssusfäden, die mit ihrem proximalen Ende am Stamm und, über eine Haftscheibe vermittelt, mit ihrem distalen Ende am Untergrund anhaften<sup>[122]</sup>. Die Byssusfäden selbst sind komplexe Kompositmaterialien, die aus bis zu zehn verschiedenen Proteinen sowie Lipiden und Wachsen bestehen und an deren Herstellung spezialisierte Enzyme beteiligt sind<sup>[37,43,123]</sup>.

### 1.3.1 Makroskopischer Aufbau des Byssusfadens

Die größtenteils aus Proteinen bestehenden, zwei bis fünf Zentimeter langen Byssusfäden können nach morphologischen sowie funktionalen Gesichtspunkten in vier makroskopisch unterscheidbare Bereiche eingeteilt werden. Proximal ist jeder Byssusfaden über eine ringartige Struktur mit dem Stamm verbunden. Das daran anschließende erste Drittel des Byssusfadens ist abgeflacht und weist eine wellige Oberfläche auf (im Folgenden als proximaler Teil bezeichnet), während der restliche Byssusfaden zylindrisch ist und eine glatte Oberfläche besitzt (im Folgenden als distaler Teil bezeichnet). Den distalen Abschluss bildet schließlich eine als Plaque oder Haftscheibe bezeichnete Struktur, die die Haftung des Byssusfadens an einer Vielzahl verschiedener Oberflächen gewährleistet. (vgl. Abb. 1-4)<sup>[42,122]</sup>.



**Abb. 1-4: Aufbau und Struktur des Muschelbyssus.** Die 2–5 cm langen Byssusfäden entspringen am Stamm und sind über diesen mit dem Muschelgewebe verbunden. Jeder Byssusfaden lässt sich in einen proximalen und distalen Teil unterteilen. Diese Unterteilung spiegelt sich auch in der Verteilung der Proteinkomponenten wieder. Die Konzentration von preCol-P und preCol-D ändert sich in gegenläufigen Gradienten entlang der Fadenlängsachse, wobei preCol-P die höchste Konzentration im proximalen und preCol-D im distalen Teil aufweist. preCol-NG hingegen ist gleichmäßig über den Faden verteilt. Jeder Byssusfaden kann über die durch verschiedene Mfp gebildete Haftscheibe an verschiedensten Substraten anhaften. Der gesamte Byssusfaden wird von der aus Mfp-1 gebildeten Kutikula umschlossen.

Sowohl der proximale wie auch der distale Teil der Byssusfäden kann, im Querschnitt betrachtet, in einen Fadenkern und eine diesen ummantelnde Kutikula unterteilt werden<sup>[37,61,62]</sup> (Abb. 1-2 C). Der Fadenkern setzt sich aus einer proteinbasierten Matrix sowie den in diese eingebetteten, smektisch angeordneten 6+1 Bündeln der Homotrimere der drei kollagenähnlichen Proteine preCol-P, -D und -NG zusammen<sup>[31,56-60]</sup>. Die sich von einem dehnbaren und elastischen proximalen zu einem steifen distalen Teil graduell entlang der Längsachse verändernden mechanischen Eigenschaften der Byssusfäden resultieren dabei in erster Linie aus den mechanischen Eigenschaften der preCols, deren spezifischer Verteilung sowie deren Quervernetzung<sup>[31,105-110]</sup>.

### 1.3.2 Die preCols als zentrale Proteinkomponente der Byssusfäden

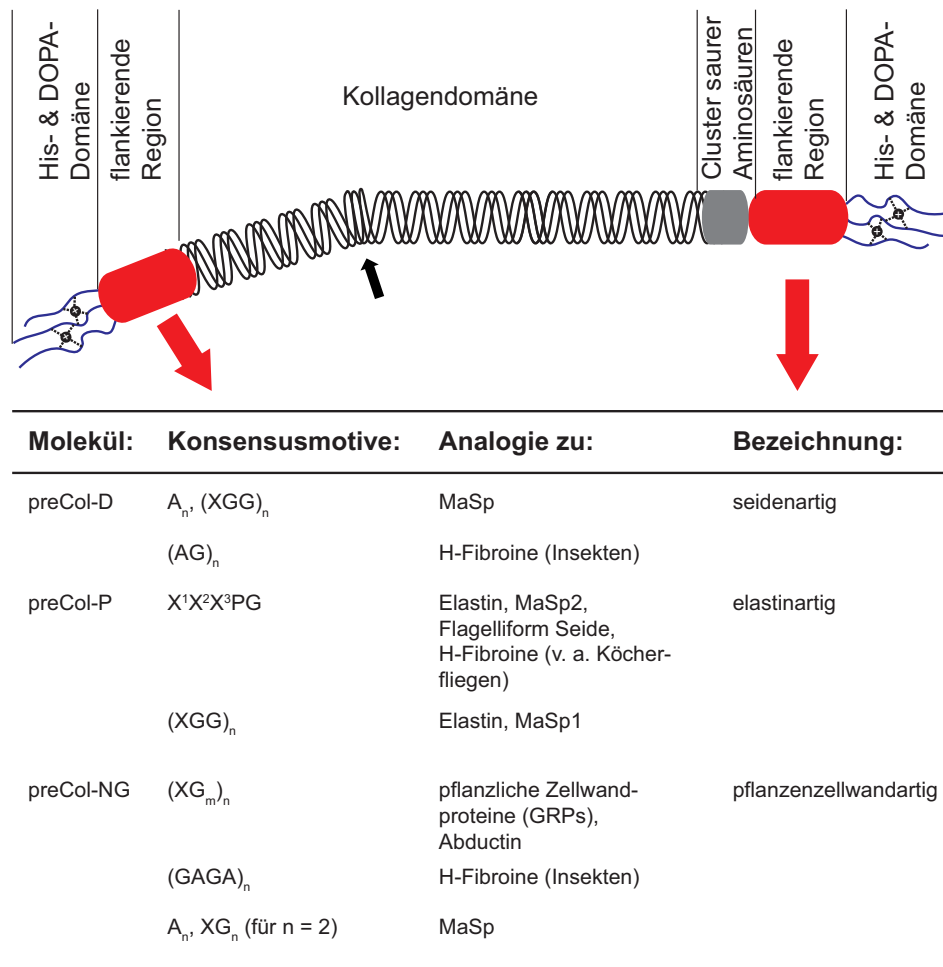
Der Kern von Byssusfäden entsteht durch die smektische (flüssigkristalline) Anordnung dreier kollagenähnlicher Proteine<sup>[31,56-58]</sup>. Zwei der drei als preCols bezeichneten (von engl. „*pre-pepsinized collagen*“) Proteine, preCol-D (distales preCol) und preCol-P (proximales preCol), sind nicht homogen entlang der Längsachse des Byssusfadens verteilt. Die höchste Konzentration an preCol-D findet sich im distalen Teil des Fadens und nimmt in proximaler Richtung kontinuierlich ab. PreCol-P weist ebenfalls einen Konzentrationsgradienten mit der höchsten preCol-P Konzentration im proximalen Teil des Fadens auf. Anders als im Fall von preCol-D erstreckt sich der preCol-P Gradient jedoch nicht über die gesamte Länge des Fadens, sondern reicht vom proximalen Ende bis etwa in die Fadenmitte<sup>[28,31]</sup>. Die Konzentration von preCol-NG (nichtgraduell verteiltes preCol, von engl. „*non-gradient*“) ist hingegen, unabhängig von der Position im Faden, weitgehend konstant<sup>[31]</sup>.

#### 1.3.2.1 Modularer Aufbau

Auf molekularer Ebene zeigen die preCols einen modularen Aufbau (Abb. 1-5). Die Kollagendomäne macht ~ 50 % der jeweiligen preCol Sequenz aus und umfasst, je nach preCol und Muschelart, zwischen 140 und 180 Wiederholungen des charakteristischen Aminosäuretripletts  $GX^1X^2$ . Wie auch bei den Wirbeltierkollagenen wird häufig die erste variable Aminosäureposition ( $X^1$ ) durch Prolin und die zweite ( $X^2$ ) durch



4-trans-Hydroxyprolin besetzt<sup>[31,56-58,93]</sup>. Bedingt durch die Kollagendomäne bilden die preCols vermutlich eine der rechtsgängigen Kollagentripelhelix ähnliche Quartärstruktur, bestehend aus drei linksgängigen,  $3_1$ -helikalen preCol-Molekülen, aus. Die entstehenden preCol-Trimere sind bereits früh während der Proteinsekretion innerhalb der sekretorischen *Granulae* nachweisbar<sup>[43,59,104]</sup>.



**Abb. 1-5: Schematischer Aufbau eines preCol-Moleküls.** Die zentrale Kollagendomäne besteht aus Wiederholungen des Tripeptidmotivs  $GX^1X^2$  und bildet eine Tripelhelix aus. Abweichungen innerhalb der  $GX^1X^2$  Sequenz führen zu lokalen Störungen der Konformation und Knicken innerhalb der Tripelhelix (schwarzer Pfeil). Die Kollagendomäne (einschließlich des *downstream* liegenden Clusters saurer Aminosäuren (grau)) wird auf beiden Seiten von flankierenden Regionen (rot) umschlossen, die je nach preCol-Typ Wiederholungen charakteristischer Konsensusmotive enthalten. Die nichtrepetitiven Termini beherbergen His-Cluster sowie DOPA-Reste, über die nichtkovalente sowie kovalente Interaktionen mit weiteren preCol-Molekülen stattfinden können. Hochgestellte Zahlen: Nummerierung der Positionen, an denen verschiedene Aminosäuren vorkommen können.

Vorhandene Abweichungen von der strikten  $GX^1X^2$ -Sequenzabfolge (Deletionen bzw. Aminosäuresubstitutionen innerhalb einzelner Triplets) führen jedoch zu lokalen

Störungen der tripelhelikalen Konformation<sup>[31,56-58]</sup>, und die Helix zeigt in diesen Bereichen eine partielle Entwindung, die sehr wahrscheinlich zu Knicken in den Homotrimeren sowie den übergeordneten 6+1 Tripelhelixbündeln führt<sup>[59,104,124-126]</sup> (Abb. 1-5, schwarzer Pfeil). Die die Kollagendomäne auf beiden Seiten umschließenden flankierenden Regionen bestehen aus Wiederholungen von Konsensusmotiven, die sehr charakteristisch für andere prominente Strukturproteine wie Seiden, Elastin oder glycinreiche Proteine aus pflanzlichen Zellwänden sind (Abb. 1-5), und stellen das am deutlichsten ausgeprägte Unterscheidungsmerkmal zwischen den drei preCol-Molekülen dar<sup>[31,56-58]</sup>. Die nichtrepetitiven Termini aller preCols sind durch das Vorkommen histidinreicher Motive und der durch posttranslationale Modifikation von Tyrosinen entstandenen DOPA-Reste gekennzeichnet<sup>[31,56-58]</sup>. Durch die Bildung von Histidin-Metall-Koordinationskomplexen mit Übergangsmetallkationen<sup>[93-98]</sup> und womöglich kovalenten di-DOPA Bindungen<sup>[99,100]</sup> spielen diese Sequenzen eine wichtige Rolle bei der Quervernetzung der preCols (Abb. 1-5).

### 1.3.2.2 Aufbau und Funktion der flankierenden Domänen

Die flankierenden Regionen von preCol-D umfassen die Sequenzmotive (A)<sub>n</sub>, (AG)<sub>n</sub> und (XGG)<sub>n</sub><sup>[31,57]</sup>. Diese Sequenzmotive sind v. a. in den repetitiven Kerndomänen von Arthropodenseidenproteinen weit verbreitet (Tab. 1-1), weswegen die entsprechenden flankierenden Regionen auch als seidenartig bezeichnet werden<sup>[20,64,65]</sup> (vgl. Abb. 1-5). Die (A)<sub>n</sub>-Sequenzen weisen jedoch, verglichen mit denen der MaSp, Substitutionen durch Aminosäuren mit großen, hydrophoben Seitenketten auf<sup>[31,57]</sup>. Nichtsdestotrotz sind auch diese Sequenzen in der Lage, eine antiparallele  $\beta$ -Faltblattkonformation einzunehmen und sich kristallin anzuordnen<sup>[73]</sup>. Gleiches gilt vermutlich für die ebenfalls vorkommenden (AG)<sub>n</sub>-Sequenzen, da diese den (XG)<sub>n</sub>-Sequenzen (X = A, V, Y) der Schmetterlings-H-Fibroine sehr ähnlich sind (vgl. Tab. 1-1). Das dritte auftretende Konsensusmotiv, (XGG)<sub>n</sub>, bildet hingegen  $\beta$ -Schleifen bzw.  $3_1$ -Helices<sup>[57,58,73,77]</sup>. V. a. die kristallin angeordneten  $\beta$ -Faltblätter sind, neben der Kollagendomäne, für die auf makroskopischer Ebene beobachtbare Steifigkeit und Zugfestigkeit distaler Byssusfäden verantwortlich<sup>[105-110]</sup>. Zudem wird vermutet, dass die stabile  $\beta$ -Faltblattkonformation der flankierenden Regionen von preCol-D dazu beiträgt, auf Byssusfäden wirkende

mechanische Kräfte zuerst auf die nichtkovalenten Histidin-Metall-Koordinationskomplexe zu übertragen (vgl. Kapitel 1.3.2.1)<sup>[94,111]</sup>. Durch diesen Mechanismus könnte mehr mechanische Energie aufgenommen werden, bevor es zu einem mit dem Brechen kovalenter Bindungen innerhalb des Proteinrückgrats einhergehendem mechanischen Totalversagen des distalen Byssus kommt.

Die glycinreichen flankierenden Regionen von preCol-P bestehen ebenfalls aus Wiederholungen von Konsensusmotiven, die aus anderen Strukturproteinen bekannt sind. Sequenzwiederholungen verschiedener Varianten des Pentapeptidmotivs  $X^1X^2X^3PG$  (wobei auf den Positionen  $X^1-X^3$  mindestens noch zwei weitere Glycine vorkommen) und  $(XGG)_n$ -Motive führen zu einer großen Ähnlichkeit mit Elastin sowie verschiedenen Seidenproteinen<sup>[31,56]</sup>. Genau wie die  $(VGVP)_n$ - und  $(XGG)_n$ -Konsensusmotive des Elastins<sup>[127]</sup> oder die  $(X^1X^2GPG)_n$ - und  $(XGG)_n$ -Wiederholungen verschiedener Arthropodenseidenproteine<sup>[3,14,20,26,27,64,65]</sup> (Tab. 1-1) falten vermutlich auch die erwähnten Konsensusmotive der flankierenden Regionen von preCol-P in  $\beta$ -Schleifen bzw.  $3_1$ -helikale Konformationen<sup>[31,56,77]</sup>. Diese Struktur motive verleihen dem proximalen Teil von Byssusfäden seine Elastizität und Dehnbarkeit<sup>[31,56]</sup>. Die mit einer Dehnung verbundene Einschränkung von Freiheitsgraden innerhalb dieser relativ flexiblen Bereiche führt zudem zu einer Entropieabnahme, die zum Aufbau einer der Dehnung entgegen wirkenden Rückstellkraft führt (sog. *elastic recoil*)<sup>[128,129]</sup>.

Die flankierenden Regionen von preCol-NG sind bezüglich ihrer repetitiven Konsensusmotive heterogener als diejenigen von preCol-D und -P. Verglichen mit letzteren unterscheiden sich amino- und carboxyterminal flankierende Region zudem deutlicher voneinander. Es sind sowohl  $(A)_n$ - wie auch  $(XG)_n$ -Sequenzen (wobei X häufig A ist)<sup>[31,58]</sup>, welche innerhalb der Arthropodenseidenproteine für die Bildung kristalliner  $\beta$ -Faltblätter verantwortlich sind<sup>[3,20,63-65]</sup>, vorhanden. Die Polyalaninsequenzen sind jedoch ausschließlich auf die aminoterminal flankierende Region beschränkt und an jeder vierten Position durch Aminosäuren mit großen Seitenketten unterbrochen<sup>[31,58]</sup>. In beiden flankierenden Regionen umschließen die seidenartigen Sequenzmotive glycinreiche, hauptsächlich aus  $XG_n$ -Wiederholungen bestehende Konsensusmotive, die aufgrund ihrer Ähnlichkeit zu in pflanzlichen Zellwänden vorkommenden glycinreichen Proteinen (GRPs, von engl. „*glycin-rich proteins*“)<sup>[130-132]</sup> namensgebend für die flankierenden Regionen des preCol-NG waren. Zudem erinnern die genannten  $XG_n$ -Motive der

flankierenden Regionen von preCol-NG auch an die aus den MaSp, H-Fibroinen und dem Elastin bekannten (XGG)<sub>n</sub>-Konsensusmotive (vgl. Tab. 1-1). Somit kombinieren sie die Primärstrukturcharakteristika der flankierenden Regionen von preCol-D und -P.

Verglichen mit der zentralen Kollagendomäne sowie den His/DOPA-Domänen ist über die Rolle und Funktion der flankierenden Regionen allerdings erst wenig bekannt. Bisherige wissenschaftliche Ergebnisse zeigen, dass sie in eine Vielzahl von Aufgaben, angefangen von der Mitbestimmung der mechanischen Eigenschaften des Byssus<sup>[105-110]</sup> sowie der Feinabstimmung der Byssusmechanik entlang der Fadenachse<sup>[31]</sup>, über die Dissipation mechanischer Energie bis hin zu Selbstheilungsprozessen des distalen Byssusteils<sup>[94,111,133]</sup>, involviert zu sein scheinen. Um die Beteiligung und Rolle der flankierenden Regionen bei diesen Prozessen jedoch detaillierter verstehen und um noch weitere Funktionen aufklären zu können, ist eine genau Aufklärung der Struktur-Funktions-Beziehungen für jede der sechs flankierenden Regionen von besonderem Interesse.

### 1.3.3 Isolierung nativer und Herstellung rekombinanter preCols

Anders als bei Spinnen- und Insektenseidenproteinen ist eine Extraktion nativer preCol-Moleküle direkt aus Byssusfäden bisher noch nicht gelungen. MaSp können als Hauptproteinbestandteil des Abseilfadens, der z. B. durch das erzwungenes „Melken“ Echter Radnetzspinnen erhalten werden kann, direkt aus diesem isoliert werden<sup>[134,135]</sup>. Auch für die Gewinnung von Fibroinen aus (*B. mori*) Kokonseide sind verschiedene chemische und enzymatische Verfahren, mit denen die Fibroine von der sie umgebenden Sericinschicht getrennt werden können, beschrieben<sup>[136-138]</sup>.

Durch die starke Quervernetzung der preCols über Histidin-Metall-Koordinationskomplexe und kovalente di-DOPA Bindungen (vgl. Kapitel 1.2.2.2) ist die Extraktion nativer preCol-Moleküle aus Byssusfäden bisher nicht gelungen. Es existieren lediglich chemische Extraktionsverfahren, mit deren Hilfe nach enzymatischer Abspaltung der Termini (einschließlich der flankierenden Regionen) nur die verbleibenden Kollagendomänen aus Byssusfäden isoliert werden können<sup>[28,31,58]</sup>. Zudem konnten Analoga nativer preCols im Gegensatz zu MaSp<sup>[139]</sup> sowie Proteinen, welche charakteristische Sequenzen von Insektenfibroinen beinhalten<sup>[140-142]</sup>, bisher auch nicht rekombinant produziert werden<sup>[36]</sup>. Dies hängt hauptsächlich damit zusammen, dass nur wenige

biotechnologisch etablierte Expressionssysteme *per se* in der Lage sind, die sowohl in der Kollagendomäne als auch den nichtrepetitiven Termini auftretenden posttranslationalen Modifikationen (v. a. Hydroxylierungen von Prolin- und Tyrosinresten) durchzuführen. Dies macht während der rekombinanten Produktion nativer preCols stets die aufwändige Coexpression funktionsfähiger Enzyme wie der Prolyl-4-hydroxylase (P4H) und einer (womöglich muschelspezifischen) Tyrosinase notwendig<sup>[143-146]</sup>, was die erfolgreiche rekombinante Produktion der preCols stark erschwert.

## 1.4 Zielsetzung

Der Muschelbyssus ist eine aus vielen Komponenten bestehende, extrakorporale Proteinstruktur. Um verstehen zu können, worauf seine außergewöhnlichen Eigenschaften beruhen, ist neben der Charakterisierung des makroskopischen Aufbaus sowie der Identifizierung aller beteiligten Moleküle v. a. auch die Identifizierung von Struktur-Funktions-Beziehungen innerhalb derselben sowie auch innerhalb einzelner Proteindomänen und Konsensusmotive nötig.

Die drei kollagenartigen Proteine preCol-D, preCol-P und preCol-NG, die den lasttragenden Kern der Byssusfäden bilden, zeigen einen modularen (domänenartigen) Aufbau ihrer Primärstruktur. Besonders auffallend sind die die zentrale Kollagendomäne flankierenden Regionen, da diese sowohl die größten Sequenzunterschiede zwischen den drei preCols aufweisen als auch Wiederholungen von Konsensusmotiven, die sonst für Seidenproteine, elastomere Proteine (wie Elastin oder Resilin) und Proteine, die in pflanzlichen Zellwänden gefunden werden, charakteristisch sind, umfassen. Bis auf ihre Aminosäuresequenz ist über die flankierenden Regionen an sich jedoch bisher relativ wenig bekannt.

Ziel der hier durchgeführten Arbeiten war es deshalb, durch Charakterisierung der flankierenden Regionen der preCols mehr über deren Aufgaben und Funktionen zu erfahren. Da sowohl eine direkte Extraktion nativer preCol-Moleküle aus Byssusfäden bisher unmöglich als auch deren rekombinante Herstellung extrem aufwändig ist, sollten ausgehend von einer cDNA-Bank des Fußes der Muschel *Mytilus galloprovincialis* nur die flankierenden Regionen (die *in vivo* keine posttranslationalen Modifikationen zeigen), ohne Kollagendomäne und nichtrepetitive Termini, rekombinant in *E. coli* hergestellt und

## 1 Einleitung

---

anschließend gereinigt werden. Durch biochemische sowie biophysikalische Analysen dieser den flankierenden Regionen nachempfundenen *flank*-Proteine sollten neue Ergebnisse hinsichtlich der Struktur-Funktions-Beziehungen und der Rolle dieser interessanten Domänen erhalten werden.

## 2 Material und Methoden

### 2.1 Verwendete Materialien

#### 2.1.1 Chemikalien

Bezeichnung:	Herkunft:
1,2-Dilauroyl- <i>sn</i> -glycero-3-phospho-(1'- <i>rac</i> -glycerol) (Natriumsalz) (DLPG)	Avanti® Polar Lipids, Alabaster, USA
1,2-Dimyristoyl- <i>sn</i> -glycero-3-phospho-(1'- <i>rac</i> -glycerol) (Natriumsalz) (DMPG)	Avanti® Polar Lipids, Alabaster, USA
1,2-Distearoyl- <i>sn</i> -glycero-3-phospho-(1'- <i>rac</i> -glycerol) (Natriumsalz) (DSPG)	Avanti® Polar Lipids, Alabaster, USA
1,2-Distearoyl- <i>sn</i> -glycero-3-phosphocholin (DSPC)	Avanti® Polar Lipids, Alabaster, USA
1,4-Dithiothreitol (DTT) ≥ 99 %, p.a.	Carl Roth, Karlsruhe
2'-Desoxyadenosin-5'-triphosphat (dATP)	New England Biolabs, Ipswich, USA
2'-Desoxycytidin-5'-triphosphat (dCTP)	New England Biolabs, Ipswich, USA
2'-Desoxyguanosin-5'-triphosphat (dGTP)	New England Biolabs, Ipswich, USA
2'-Desoxythymidin-5'-triphosphat (dTTP)	New England Biolabs, Ipswich, USA
2-Amino-2-(hydroxymethyl)-propan-1,3-diol (Tris)	Carl Roth, Karlsruhe
2-Mercaptoethanol	Carl Roth, Karlsruhe
3-(N-Morpholino)-propansulfonsäure (MOPS, Handelsname: PUFFERAN®) ≥ 99,5 %	Carl Roth, Karlsruhe
Agar-Agar	Carl Roth, Karlsruhe
Agarose	Carl Roth, Karlsruhe
Ammoniak 25 %	Carl Roth, Karlsruhe
Ammoniumhydrogencarbonat	Carl Roth, Karlsruhe
Ammoniumperoxodisulfat (APS)	Carl Roth, Karlsruhe
Ammoniumsulfat	Carl Roth, Karlsruhe
Ampicillin Natriumsalz	Carl Roth, Karlsruhe
Breox FMT 30 Antifoam	Cognis, Southampton, Großbritannien
Bromphenolblau	Carl Roth, Karlsruhe
Calciumchlorid-Dihydrat	Carl Roth, Karlsruhe
Chloramphenicol	Carl Roth, Karlsruhe
Coomassie Brilliant Blau	Carl Roth, Karlsruhe

## 2 Material und Methoden

---

di-Kaliumhydrogenphosphat	Carl Roth, Karlsruhe
di-Natriumhydrogenphosphat	Carl Roth, Karlsruhe
Essigsäure	VWR International, Radnor, USA
Ethanol (technisch und p. A.)	VWR International, Radnor, USA
Ethylendiamintetraessigsäure (EDTA)	Carl Roth, Karlsruhe
Formaldehyd 37 %	Carl Roth, Karlsruhe
Glucose	Carl Roth, Karlsruhe
Glycerin 99,5 %	Carl Roth, Karlsruhe
Glycin	Carl Roth, Karlsruhe
Guanidin-hydrochlorid (GuaHCl) $\geq 99,7$ %, p. a., protein grade	Carl Roth, Karlsruhe
Guanidinthiocyanat (GuaSCN) $> 99$ %, reinst	Carl Roth, Karlsruhe
Harnstoff	Carl Roth, Karlsruhe
Hefeextrakt	Carl Roth, Karlsruhe
Imidazol	Sigma-Aldrich, St. Louis, USA
Isopropanol	VWR International, Radnor, USA
Isopropyl-1-thio- $\beta$ -D-galactopyranosid (IPTG)	Carl Roth, Karlsruhe
Kaliumdihydrogenphosphat	Carl Roth, Karlsruhe
Kaliumsulfat	Carl Roth, Karlsruhe
Kanamycin-Sulfat	Carl Roth, Karlsruhe
LB-Agar	Carl Roth, Karlsruhe
LB-Medium	Carl Roth, Karlsruhe
Magnesiumchlorid-Hexahydrat	Carl Roth, Karlsruhe
Magnesiumsulfat-Heptahydrat	Carl Roth, Karlsruhe
Mangan(II)-chlorid-Tetrahydrat	Carl Roth, Karlsruhe
Methanol	VWR International, Radnor, USA
Milchpulver Blotting Grade	Carl Roth, Karlsruhe
N,N,N',N'-Tetramethylethylendiamin (TEMED)	Carl Roth, Karlsruhe
Natriumacetat	Carl Roth, Karlsruhe
Natriumcarbonat	Carl Roth, Karlsruhe
Natriumchlorid	Carl Roth, Karlsruhe
Natriumdodecylsulfat (SDS)	Carl Roth, Karlsruhe
Natriumhydroxid	Carl Roth, Karlsruhe
Natriumthiosulfat-Pentahydrat	Carl Roth, Karlsruhe
Phenylmethylsulfonylfluorid (PMSF)	Carl Roth, Karlsruhe
Proteaseinhibitor Mix G	Serva Electrophoresis, Heidelberg
Rotiphorese Gel 40 (37,5 : 1)	Carl Roth, Karlsruhe



Rotiphorese Gel 40 (19 : 1)	Carl Roth, Karlsruhe
Rotiphorese Gel 40 (29 : 1)	Carl Roth, Karlsruhe
Salzsäure 37 %	Carl Roth, Karlsruhe
Sauerstoff	Rießner-Gase, Lichtenfels
Silbernitrat	Carl Roth, Karlsruhe
Stickstoff (flüssig)	Linde, München
SYBR Safe DNA Gel Stain	Invitrogen (Life Technologies), Carlsbad, USA
Tetracyclin	Carl Roth, Karlsruhe
Trichlormethan (Handelsname: ROTIPURAN®), ≥ 99 %, p. a.	Carl Roth, Karlsruhe
Trypton/Pepton aus Casein	Carl Roth, Karlsruhe
Xylencyanol	Carl Roth, Karlsruhe

### 2.1.2 Größenstandards und Kits

<b>Bezeichnung:</b>	<b>Herkunft:</b>
100 bp DNA-Leiter	New England Biolabs, Ipswich, USA
1 kb DNA-Leiter	New England Biolabs, Ipswich, USA
Amersham ECL Plus Western Blotting Detection System	GE Healthcare, Chalfont St. Giles, Großbritannien
<i>ColorPlus™ Prestained Protein Marker, Broad Range (7–175 kDa)</i>	New England Biolabs, Ipswich, USA
<i>Wizard® Plus SV Minipreps DNA Purification System</i>	Promega, Fitchburg, USA
<i>Wizard® SV Gel and PCR Clean-Up System</i>	Promega, Fitchburg, USA

### 2.1.3 Oligonukleotide und artifizielle Nukleotidsequenzen

#### Oligonukleotide

Oligonukleotide wurden als Primer zur Amplifikation von *M. galloprovincialis* preCol-Nukleotidsequenzen mittels PCR, als Sequenzierprimer zur Überprüfung der erzeugten Vektorkonstrukte und zur Mutagenese von Nukleotidsequenzen verwendet. Die Oligonukleotide wurden von den Firmen Eurofins MWG Operon, biomers.net und Microsynth AG bezogen. Zusätzliche Sequenzierprimer waren Standardprimer der die DNA-Sequenzierungen durchführenden Firma GATC Biotech und wurden von dieser direkt vor Ort synthetisiert und verwendet.

## 2 Material und Methoden

**Tab. 2-1. Übersicht der verwendeten Oligonukleotide.** In den Sequenzen wurden Erkennungs- und Schnittsequenzen für Restriktionsendonukleasen grün und durch die Oligonukleotide zusätzlich eingeführte Sequenzabschnitte blau markiert. Bei den als Mutagenese-Primer verwendeten Oligonukleotiden wurden von der Originalsequenz abweichende Basen rot markiert. fwd bzw. f: Vorwärtsprimer; rev: Rückwärtsprimer; I: Eurofins MWG Operon, Ebersberg; II: biomers.net GmbH, Ulm; III: Microsynth AG, Balgach, Schweiz; IV: GATC Biotech, Konstanz.

Bezeichnung:	Sequenz (5' nach 3'):	Verwendung:	Herkunft:
PreColP-ntf_fwd	AGCGCCAACGCTCAC ATATGGGTGGATTCTG GTGG	Amplifikation der 5'-flankierenden Region von preCol-P aus cDNA-Bank Mfl-1;	I
PreColP-ntf_rev	AGTCGAATTCTCATT AGCCGAATCCTCCAC CACCATTGTCATTTG	PCR-Screening nach Ligation	
PreColP-3'_Mg_fwd	GACAGTACATATGTT CCCAGGATTTGGTGG ATTTGGTGGTTTC	Amplifikation der 3'-flankierenden Region von preCol-P aus cDNA-Bank Mfl-1;	I
PreColP-ctf_rev	GCGCGAATTCTCATT ATGCTCTAGATGCTG CAGATGCTGATGCGC	PCR-Screening nach Ligation	
PreColD-5'_fwd	TTCCCACATATGGGT GGTAGCGCAAGCGCA GC	Amplifikation der 5'-flankierenden Region von preCol-D aus cDNA-Bank Mfl-1;	I
PreColD-5'_n_rev	ACCTGGGAATTCTCA TTAGACTCCTCCGAA TCCTCC	PCR-Screening nach Ligation	
PreColD-3'_fwd	AGGAGCACATATGGG TGGACTCGGAGGAGC	Amplifikation der 3'-flankierenden Region von preCol-D aus cDNA-Bank Mfl-1;	I
PreColD-3'_rev	CTCGTGTGAATTCTC ATTATTGTGCGGCTG CGTTGGC	PCR-Screening nach Ligation	
preColNG-ntf_fwd	TCATGGAGGAAGTGC CATATGGCAGCAGCC CATGCAG	Amplifikation der 5'-flankierenden Region von preCol-NG aus cDNA-Bank Mfl-1;	I
preColNG-ntf_rev	AGTCGAATTCTCATT AACCTGGTAAGGGGA ATCCTCCTCCGCCAG	PCR-Screening nach Ligation	
preColNG-ctf_fwd	GTGGACCAGCTGGCA TATGGGTGCAGGAGA AGGCGG	Amplifikation der 3'-flankierenden Region von preCol-NG aus cDNA-Bank Mfl-1	I
preColNG-ctf_rev	AGTCGAATTCTCATT AAGCACTGGCACTAC CGCCTGCACCTC		

Bezeichnung:	Sequenz (5' nach 3'):	Verwendung:	Herkunft:
eCTSUMO_fwd	ACAACA <b>GGATCCTAT</b> CCAGGATTTGGTGA TTTGGTG	Amplifikation der eCT-Sequenz aus pET28a-eCT, zusätzlich: fwd: Phe- → Tyr-Mutation;	II
eCTSUMO_rev	ACAACA <b>AAGCTT</b> TCA TTA <b>CGCATAT</b> GTCTCT AGATGCTGCAG	rev: Sequenzverlängerung um His- & Ala-Codon; PCR-Screening nach Ligation	
cwNTSUMO_fwd <sup>1</sup>	CTACA <b>GGATCCTAT</b> G CAGCAGCCCATGCAG CAG	Amplifikation der cwNT-Sequenz aus pET28a-cwNT; zusätzlich: fwd: Sequenzverlängerung um Tyr-Codon;	II
cwNTSUMO_rev <sup>1</sup>	GGACA <b>AAGCTT</b> TTATC <b>AGGCATA</b> ACCTGGTA AGGGGAATCCTCCTC CG	rev: Sequenzverlängerung um Tyr- & Ala-Codon	
cwCTSUMO_fwd <sup>1</sup>	CTACA <b>GGATCCTAT</b> G GTGCAGGAGAAGGCG GAG	Amplifikation der cwCT-Sequenz aus pET28a-cwNT; zusätzlich: fwd: Sequenzverlängerung um Tyr-Codon;	II
cwCTSUMO_rev <sup>1</sup>	GGACA <b>AAGCTT</b> TTATC <b>AGGCATA</b> AGCACTGG CACTACCGCC	rev: Sequenzverlängerung um Tyr- & Ala-Codon	
Sumo_cwCT(F34W)_f	CGAATC <b>GGATCC</b> ATG GGTGCCGGTGAAGGT GGTG	Amplifikation der cwCT(F34W)-CBD-Sequenz aus pTXB1-cwCT(F34W)	III
Sumo_cwCT(F34W)_Int_rev	CGAATC <b>AAGCTT</b> TTA TCATTGAAGCTGCCA CAAGGC		
eCTBamHI <sub>mut2</sub> _fwd	GCTGGTGGACCCGGT GG <b>T</b> TCCGCTGGAGCA GGTTC	Mutagenese von pET28a-eCT zur Entfernung der BamHI-Erkennungssequenz in eCT	II
eCTBamHI <sub>mut2</sub> _rev	GAACCTGCTCCAGCG GA <b>A</b> CCACCGGGTCCA CCAGC		
T7	TAATACGACTCACTA TAGGG	Sequenzierprimer	IV
pET-RP	CTAGTTATTGCTCAG CGG	Sequenzierprimer	IV
SUMO-seq	CGGTATTAGAATTCA AGCTGATCAGACCCC	Sequenzierprimer	IV
GATC-intein_rev-508922	CAAGGACTTTCAGGT CGATGGCG	Sequenzierprimer	IV

<sup>1</sup> Von Martina B. Elsner im Rahmen ihrer Masterarbeit <sup>[147]</sup> entworfen und verwendet

### Artifizielle Nukleotidsequenzen:

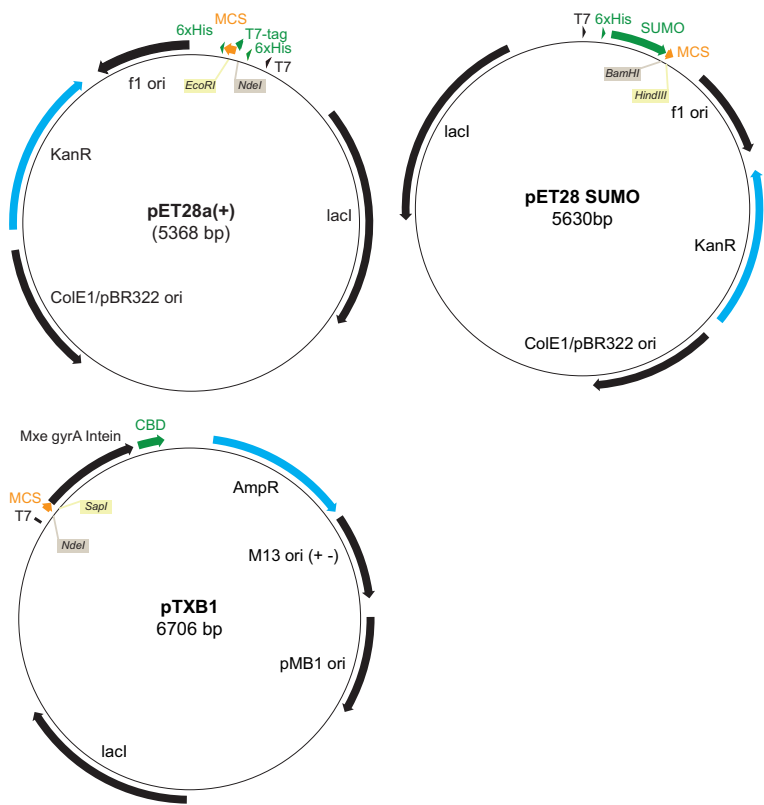
Die für die Erzeugung der eCT(F106W)- und cwCT(F34W)-Vektorkonstrukte (Tab. 2-3) verwendeten Nukleotidsequenzen wurden von der Firma GeneArt (inzwischen: Life Technologies) an die Codonverwendung von *E. coli* angepasst, synthetisiert und als in Plasmide ligierte, doppelsträngige Nukleotidsequenzen geliefert.

**Tab. 2-2. Artifizielle Nukleotidsequenzen.** Die Erkennungs- und Schnittsequenzen für Restriktionsendonukleasen wurden grün markiert und die für das jeweilige Protein codierende Sequenz unterstrichen.

Bezeichnung:	Sequenz (5' nach 3'):
eCT(F106W)_2	<p>GAGCTC<u>GGATCC</u>CCTGGTTTTGGTGGCTTTGGCGGTTTTGGAGGTGCAAGCGC</p> <p>AAATGCAGGTAGCAGCGCCAATGCAAATGCCGGTGGTCCGGGTGGTAGCGCAG</p> <p>GCGCAGGTAGCTCAAGCGGTGCAAATGCGAATGCAGGCGGTTTTCCGTTTGGT</p> <p>GGTCCGTTTTGGCGGTGTGGGTGGTGGACCTGGTGGTGCTGGCGGTCCAGGTGG</p> <p>TGTTGGAGGTGGCCCTGGGGGTGTTGGTGGCGGTGTTGGCGGAGGTCCCCGGTG</p> <p>GCGTAGGTGGTGGGGTTGGTGGTGGTCCAGGCGGTGCCGGTGGCCTGGGAGGA</p> <p>TTTGGTGGTTGGGGTGGCGGTAGCAGTGCCGGTGCAAGCAGCAGCGGTAGCGC</p> <p>CAGCGCAAGCAGTGGTGGCCCTTTTGGTGTCTGAATGTTGGCCCTGGTGGCG</p> <p>GAATCGGTGGTGGTTCAGCAAGCGCGAGCGCAGCAAGCCGTGCCTAATAA<u>AAG</u></p> <p><u>CTT</u>GGTACC</p>
cwCT(F34W)	<p>GGCGAATTGGCGGAAGGCCGTCAAGGCCTAGGCGCGCCATGAGCT<u>CATAT</u>GGG</p> <p>TGCCGGTGAAGGTGGTGCCGGTGGTGCAGGCGGTGGTGCGGGTGGTGCCGGTG</p> <p>GCCTGGGTGGCGGTGCCGGTGGTGCCGGTGGTCTGGGTGGTGGTTGGGGTGGC</p> <p>CTGGGTGGTGGTGCGGGTGGTCTGGGTGGCCTGGGTGGTGGCGCAGGCGGTGC</p> <p>CGGTGGCGTTGGTGGCCTGGGTGGTGGTGGTGGCCTGGGTGGCGTTGGTG</p> <p>GTCTGGGCTTTGGTGGTGCCGGTGCAAGTGCCGGTGCCGGTGCTAATGCCGGT</p> <p>GCTGGCGGTGCCGGTGGCAGCGCAAGTGCAATGCGGAAGAGCAAGCTTGGTACC</p> <p>TCTTAATTAAGTGGCCTCATGGGCCTTCCGCTCACTGCCC</p>

### 2.1.4 Plasmide und Vektorkonstrukte

Die Plasmide pET28a(+), pET28 SUMO und pTXB1 wurden als Vektoren für die heterologe Expression der von den cDNA-Sequenzen der flankierenden Regionen der *M. galloprovincialis* preCols abgeleiteten Nukleotidsequenzen unter Kontrolle des T7-Promotorsystems in *E. coli* verwendet.



**Abb. 2-1. Karten der verwendeten Expressionsvektoren.** Bei den verwendeten Vektoren pET28a(+), pET28 SUMO und pTXB1 handelt es sich um auf dem T7-Promotorsystem beruhende Expressionsvektoren. Farblich markiert sind die jeweilige Klonierungskassette (MCS, orange), die 5' bzw. 3' der MCS codierten tag-Sequenzen (grün), der Selektionsmarker (blau) sowie die Replikationsursprünge (ori, schwarz). Die zur Ligation transgener Sequenzen in die MCS verwendeten Restriktionsschnittstellen wurden entsprechend ihrer Lage eingezeichnet.

Tab. 2-3 gibt einen Überblick über die Herkunft der Plasmide, über die durch Ligation der mittels PCR (vgl. Kapitel 2.3.4) amplifizierten oder artifiziellen Nukleotidsequenzen (vgl. Kapitel 2.1.3) mit den Expressionsvektoren erzeugten Vektorkonstrukte sowie über die zur Ligation verwendeten Restriktionsschnittstellen.

**Tab. 2-3. Erzeugte Vektorkonstrukte.** Übersicht über die auf Grundlage der verwendeten Nukleotidsequenzen und Expressionsvektoren erzeugten Vektorkonstrukte. Zur Ligation der Nukleotidsequenz in das Plasmid wurden die angegebenen Restriktionsendonukleasen verwendet. „5'“ und „3'“ in der Spalte Besonderheiten beziehen sich auf die Lage der tag-Sequenz relativ zur Klonierungskassette (MCS).

Plasmid:	Herkunft:	Konstrukt:	kloniert mit:	Besonderheiten:
pET28a(+)	Novagen (Merck), Darmstadt	pET28a-eNT pET28a-eCT pET28a-sNT pET28a-sCT pET28a-cwNT pET28a-cwCT	NdeI EcoRI	5': (His) <sub>6</sub> -Sequenz & Thrombinschnittstelle

Plasmid:	Herkunft:	Konstrukt:	kloniert mit:	Besonderheiten:
pET28 SUMO	Lehrstuhl für Biotechnologie, TU München	pETSUMO-eCT <sub>2Y</sub> pETSUMO-cwNT <sub>2Y</sub> <sup>2</sup> pETSUMO-cwCT <sub>2Y</sub> <sup>2</sup> pETSUMO-eCT(F106W) pETSUMO-cwCT(F34W) pETSUMO-cwCT(F34W)-CBD	<i>BamHI</i> <i>HindIII</i>	5': (His) <sub>6</sub> -Epitop & SUMO-tag
pTXB1	New England Biolabs, Ipswich, USA	pTXB1-cwCT(F34W)	<i>NdeI</i> <i>SapI</i>	3': autokatalytisch spleißende MiniIntein-Sequenz & CBD-tag  Ausgangsvektor für die Klonierung von pETSUMO-cwCT(F34W)-CBD

### 2.1.5 *E. coli*-Stämme

Tab. 2-4 zeigt die bei der Erzeugung und Vervielfältigung der verwendeten Vektorkonstrukte (vgl. Tab. 2-3) sowie die zur rekombinanten Produktion von *M. galloprovincialis* flank-Proteinen zum Einsatz gekommenen *E. coli*-Stämme.

Tab. 2-4. Übersicht der verwendeten *E. coli*-Stämme

Bezeichnung:	Verwendung:	Genotyp:	Herkunft / Ursprung:
DH10B	Klonierung & Vervielfältigung von Vektorkonstrukten	F <sup>-</sup> <i>mcrA</i> Δ( <i>mrr-hsdRMS-mcrBC</i> ) Φ80/ <i>lacZ</i> Δ <i>M15</i> Δ <i>lacX74</i> <i>recA1</i> <i>endA1</i> <i>araD139</i> Δ( <i>ara, leu</i> )7697 <i>galU</i> <i>galK</i> λ <sup>-</sup> <i>rpsL</i> <i>nupG</i>	Life Technologies (Invitrogen), Carlsbad, USA
BL21(DE3)	Expressionstests	<i>E. coli</i> B <i>gal</i> <i>hsdS<sub>B</sub></i> <i>ompT</i>	Agilent Technologies Genomics (Stratagene), Santa Clara, USA

<sup>2</sup> Von Martina B. Elsner im Rahmen ihrer Masterarbeit<sup>[147]</sup> erzeugt und verwendet

Bezeichnung:	Verwendung:	Genotyp:	Herkunft / Ursprung:
BL21-CodonPlus(DE3)-RIPL	Expressionstests & rekombinante Proteinproduktion	<i>E. coli</i> B F <sup>-</sup> <i>ompT</i> <i>hsdS</i> (rB <sup>-</sup> mB <sup>-</sup> ) <i>dcm</i> <sup>+</sup> Tet <sup>r</sup> <i>gal</i> λ(DE3) <i>endA</i> Hte [ <i>argU proL Cam</i> <sup>r</sup> ] [ <i>argU ileY leuW</i> Strep/Specr]	Agilent Technologies Genomics (Stratagene), Santa Clara, USA
BL21-Gold(DE3)	Expressionstests & rekombinante Proteinproduktion	<i>E. coli</i> B F <sup>-</sup> <i>ompT</i> <i>hsdS</i> (rB <sup>-</sup> mB <sup>-</sup> ) <i>dcm</i> <sup>+</sup> Tet <sup>r</sup> <i>gal</i> λ(DE3) <i>endA</i> Hte	Agilent Technologies Genomics (Stratagene), Santa Clara, USA
BLR(DE3)	Expressionstests & rekombinante Proteinproduktion	<i>E. coli</i> B <i>gal</i> <i>hsdS</i> <sub>B</sub> <i>ompT</i> Δ( <i>srf</i> <i>recA</i> )306::Tn10(Tet <sup>r</sup> )	Agilent Technologies Genomics (Stratagene), Santa Clara, USA
ER 2566	Expressionstests	<i>fhuA2 lacZ::T7 gene1</i> [ <i>lon</i> ] <i>ompT gal</i> <i>sulA11</i> R( <i>mcr-73::miniTn10-Tet</i> <sup>S</sup> )2 [ <i>dcm</i> ] R( <i>zgb-210::Tn10--Tet</i> <sup>S</sup> ) <i>endA1</i> Δ( <i>mcrC-mrr</i> )114::IS10	New England Biolabs, Ipswich, USA
HMS174(DE3)	Expressionstests	F <sup>-</sup> <i>recA1 hsdR</i> (rK12 <sup>-</sup> mK12 <sup>+</sup> ) (DE3) (Rif <sup>r</sup> )	Novagen® (Merck Millipore), Billerica, USA

### 2.1.6 Proteine und Peptide

#### Antikörper:

Bezeichnung:	Ursprungsorganismus:	Herkunft:
Anti-CBD Monoclonal Antibody	Maus (monoklonal)	New England Biolabs, Ipswich, USA
Anti-Mouse IgG (Whole Molecule)-Peroxidase Antibody	Ziege (polyklonal)	Sigma-Aldrich, St. Louis, USA
His•Tag® Antibody HRP Conjugate	nicht genannt, (monoklonal)	Novagen® (Merck Millipore), Billerica, USA
Monoclonal Anti-polyHistidine Antibody	Maus (monoklonal)	Sigma-Aldrich, St. Louis, USA
Polyclonal Antibody to Mouse IgG [H&L] – FITC -	Ziege (polyklonal)	Acris Antibodies, Hiddenhausen

#### Ligase:

Bezeichnung:	Herkunft:
T4 DNA Ligase	Promega, Fitchburg, USA

### Restriktionsendonukleasen:

Bezeichnung:	Erkennungssequenz/Schnittposition*:	Herkunft:
<i>BamHI</i>	5' - ... <b>G↓GATC</b> C... - 3' 3' - ... <b>C CTAG↑G</b> ... - 5'	New England Biolabs, Ipswich, USA
<i>BsgI</i>	5' - ... <b>GTGCAG</b> (N) <sub>13</sub> NN↓... - 3' 3' - ... <b>CACGTC</b> (N) <sub>13</sub> ↑NN... - 5'	New England Biolabs, Ipswich, USA
<i>DpnI</i> <sup>†</sup>	5' - ... <b>GA↓TC</b> ... - 3' 3' - ... <b>CT↑AG</b> ... - 5'	New England Biolabs, Ipswich, USA
<i>EcoRI</i>	5' - ... <b>G↓AATT</b> C... - 3' 3' - ... <b>C TTAA↑G</b> ... - 3'	New England Biolabs, Ipswich, USA
<i>HindIII</i>	5' - ... <b>A↓AGCT</b> T... - 3' 3' - ... <b>T TCGA↑A</b> ... - 5'	New England Biolabs, Ipswich, USA
<i>NdeI</i>	5' - ... <b>CA↓TA</b> TG... - 3' 3' - ... <b>GT AT↑AC</b> ... - 5'	New England Biolabs, Ipswich, USA
<i>SapI</i>	5' - ... <b>GCTCTTCN↓NNN</b> ... - 3' 3' - ... <b>CGAGAAGN NNN↑</b> ... - 5'	New England Biolabs, Ipswich, USA
<i>XbaI</i>	5' - ... <b>T↓CTAG</b> A... - 3' 3' - ... <b>A GATC↑T</b> ... - 5'	New England Biolabs, Ipswich, USA
<i>XhoI</i>	5' - ... <b>C↓TCGA</b> G... - 3' 3' - ... <b>G AGCT↑C</b> ... - 5'	New England Biolabs, Ipswich, USA

\*Erkennungssequenzen wurden grün markiert, Pfeile kennzeichnen die Restriktionsstelle; N steht für ein beliebiges Nukleotid

<sup>†</sup>Restriktion erfolgt ausschließlich nach 6-N-Methyladenosinen

### Polymerasen:

Bezeichnung:	Herkunft:
<i>PfuTurbo</i> <sup>®</sup> DNA Polymerase	Stratagene, La Jolla, USA
Phusion <sup>®</sup> High-Fidelity DNA Polymerase	New England Biolabs, Ipswich, USA
<i>Taq</i> DNA Polymerase	New England Biolabs, Ipswich, USA

### Proteasen:

Bezeichnung:	Erkennungssequenz/Schnittposition:	Herkunft:
Ulp-1	amino-säuresequenzunabhängige Erkennung der SUMO-Tertiärstruktur	Lukas Eisoldt, Lehrstuhl Bio- materialien, Universität Bayreuth

### Weitere Proteine<sup>3</sup>:

Bezeichnung:	Herkunft:
bovines Serumalbumin, Fraktion V (BSA)	Carl Roth, Karlsruhe
eADF4(C16) <sup>[148]</sup>	Claudia Blüm, Lehrstuhl Bio- materialien, Universität Bayreuth

<sup>3</sup> physikalisch-chemische Eigenschaften vgl. Kapitel 11.1.4.1



Bezeichnung:	Herkunft:
Sfl <sup>[149]</sup>	Christian Ackerschott, Lehrstuhl für Biotechnologie, TU München

### 2.1.7 Chromatographiematerialien

Bezeichnung:	Chromatographiemethode:	Herkunft:
Chelating Sepharose Fast Flow	präparative Affinitätschromatographie	GE Healthcare Europe GmbH, Freiburg
HiLoad 16/600 Superdex 30 PG	präparative Gelpermeationschromatographie	GE Healthcare Europe GmbH, Freiburg
HisTrap FF Crude 5 ml	analytische & präparative Affinitätschromatographie	GE Healthcare Europe GmbH, Freiburg
Superdex Peptide 10/300 GL	analytische Gelpermeationschromatographie	GE Healthcare Europe GmbH, Freiburg

### 2.1.8 Sonstige Materialien

3000 Series Nanosphere™ Size Standards	Thermo Fisher Scientific, Waltham, USA
Bechergläser	VWR International, Radnor, USA
Dewargefäß Agil 2	Air Liquide, Paris, Frankreich
Dialyseklammern und -schläuche (MWCO: 0,5–1,0 kDa/6,0–8,0 kDa)	Spectrum Laboratories, Rancho Dominguez, USA
Elektroporationsküvetten (0,2 cm Schichtdicke)	VWR International, Radnor, USA
Erlenmeyerkolben	VWR International, Radnor, USA
Glasperlen	Carl Roth, Karlsruhe
Indikatorstreifen pH-Fix	Carl Roth, Karlsruhe
Kanülen	B. Braun Melsungen, Melsungen
Laborflaschen	Duran Group, Mainz
Labortücher Kimtech Science	Kimberly-Clark, Dallas, USA
Messzylinder	VWR International, Radnor, USA
Mikroküvetten	Sarstedt, Nümbrecht
Omnifix Spritzen (2–50 ml)	B. Braun Melsungen, Melsungen
Parafilm M	Bemis, Neenah, USA
Pasteurpipetten	VWR International, Radnor, USA
Petrischalen	Sarstedt, Nümbrecht
Pipetten (1000, 200, 20, 10, 2,5 µl)	Eppendorf, Hamburg
Pipettenspitzen	Sarstedt, Nümbrecht
Reaktionsgefäße (200 µl, 1,5 & 2 ml)	Sarstedt, Nümbrecht

## 2 Material und Methoden

---

Roti®-PVDF Blotting-Membran	Carl Roth, Karlsruhe
Rührfische	VWR International, Radnor, USA
Serologische Pipetten (2–50 ml)	Sarstedt, Nümbrecht
SERVAGE™ TG 12, Vertical Tris-Glycine Gel 12 %	Serva Electrophoresis, Heidelberg
SERVAGE™ TG 4–20, Vertical Tris-Glycine Gel 4–20 %	Serva Electrophoresis, Heidelberg
Skalpelle	B. Braun Melsungen, Melsungen
Sterilfilter Filtropur S (0,22 & 0,45 µm)	Sarstedt, Nümbrecht
Suprasil-Küvetten (Quarzglas; 1, 10 mm)	Hellma Analytics, Müllheim
TEM-Objektträger auf Kupfer, Pioloform-Beschichtung	Lehrstuhl Zellbiologie und Elektronenmikroskopie, Universität Bayreuth, Bayreuth & Plano GmbH, Wetzlar
UV-Küvetten	Brand, Wertheim
Wägeschälchen	VWR International, Radnor, USA
Whatman™ 3MM CHR Chromatographiepapier	GE Healthcare, Chalfont St. Giles, Großbritannien
Zentrifugenröhrchen (15 & 50 ml)	VWR International, Radnor, USA
ZipTip C4 & C18-Pipettenspitzen	Merck Millipore, Billerica, USA

### 2.1.9 Geräte

#### Anzucht/Inkubation:

Autoklav SANoclav	Adolf Wolf SANoclav, Bad Überkingen-Hausen
Autoklav VX-150	Systemec, Wetzlar
5 l-Glasgefäß, begasbar	Glasbläserei Universität Bayreuth, Bayreuth
Bioreaktor Minifors 1,3 l	Infors HT, Bottmingen, Schweiz
Tischinkubatoren Thermomixer comfort & Thermomixer compact	Eppendorf, Hamburg
Wärmeschrank WB560K	Mytron, Heilbad Heiligenstadt
Wasserbad Aqualine AL 12	Lauda, Lauda-Königshofen
Wasserbad F12-MB	Julabo, Seelbach
Wasserbad Haake DC30-K20	Thermo Fisher Scientific, Waltham, USA
Wasserbad Haake WKL 26	Thermo Fisher Scientific, Waltham, USA

#### Aufbewahrung:

Gefrierschrank Comfort NoFrost	Liebherr, Bulle, Schweiz
Kühl-Gefrierschrank Premium	Liebherr, Bulle, Schweiz
Tiefkühlagertruhe HERAFreeze	Thermo Fisher Scientific, Waltham, USA

Gefriertrockner Alpha 1-2 LDplus	Christ, Osterode am Harz
<b><u>Bildgebung/Dokumentation:</u></b>	
Digitale Spiegelreflexkamera D5000	Nikon, Tokyo, Japan
Objektiv AF 90 mm F/2.8 Macro 1:1	Tamron, Saitama, Japan
Fluoreszenzbildgeber Ettan DIGE	GE Healthcare, Chalfont St. Giles, Großbritannien
Transilluminator Geldoc XR	Bio-Rad, Hercules, USA
Gelscanner Bio-5000	Microtek, Hsinchu, Taiwan
Gelscanner HP Scanjet G4050	Hewlett-Packard, Palo Alto, USA
<b><u>Chromatographie:</u></b>	
Äkta FPLC	GE Healthcare, Chalfont St. Giles, Großbritannien
Äkta Purifier	GE Healthcare, Chalfont St. Giles, Großbritannien
Fraktionssammler Frac-950	GE Healthcare, Chalfont St. Giles, Großbritannien
Probenauftragspumpe P-960 mit Luftsensoren	GE Healthcare, Chalfont St. Giles, Großbritannien
<b><u>Elektrophorese:</u></b>	
Agarose-Gelelektrophoresekammern: BlueMarine™ 100 & BlueMarine™ 200	Serva Electrophoresis, Heidelberg
Heizblock	VWR International, Radnor, USA
Fastblot	Biometra, Jena
PAGE-Elektrophoresekammer Mighty Small II	Hoefer, Holliston, USA
Polyacrylamidgel-Gießstand	Mechanikwerkstatt Universität Bayreuth, Bayreuth
Spannungsquellen: BluePower™ 500 & BluePower™ Plus	Serva Electrophoresis, Heidelberg
<b><u>Herstellung von Medien/Puffern/Lösungen:</u></b>	
Feinwaage TE313S	Sartorius, Göttingen
pH-Meter Schott Lab 850	SI Analytics, Mainz
Reinstwasseranlage MilliQ Advantage A10	Merck Millipore, Darmstadt
Waage Acculab ALC-4100.1	Sartorius, Göttingen
Waage CPA225D	Sartorius, Göttingen
Waage CS200	Ohaus, Pine Brook, USA

Waage Pioneer	Ohaus, Pine Brook, USA
<b><u>Mischen/Homogenisieren:</u></b>	
Dispergierer Silent Crusher M	Heidolph, Schwabach
Laborrührer RZR 2020	Heidolph, Schwabach
Magnetrührer KMO 2 basic	IKA, Staufen
Magnetrührer mit Heizplatte MR 3001	Heidolph, Schwabach
Magnetrührer mit Heizplatte RTC basic	IKA, Staufen
Magnetrührer VS-C7	VWR International, Radnor, USA
Schüttler Certomat RM	Sartorius, Göttingen
Schüttler Unimax 1010	Heidolph, Schwabach
Taumelschüttler 3011	GFL, Burgwedel
Überkopfschüttler Intelli-Mixer	neoLab, Heidelberg
Ultraschalleinheit Sonopuls HD 3200	Bandelin electronic, Berlin
Ultraschallsonden KE 76, MS 73 & TT 13	
Vortex VV3	VWR International, Radnor, USA
<b><u>Spektroskopie:</u></b>	
Jasco FP-6500 Spektrofluorimeter	Jasco, Tokyo, Japan
temperierbare Messeinheit ADP-303	
Pelletierelement ADP-303T	
Jasco J-815 Spektropolarimeter	Jasco, Tokyo, Japan
Pelletierelement PTC 423S/15	
MALDI-TOF-Massenspektrometer Reflex II	Bruker, Billerica, USA
NanoDrop ND-100 Photometer	Thermo Fisher Scientific, Waltham, USA
OD600 DiluPhotometer	Thermo Fisher Scientific, Waltham, USA
UV/VIS-Spektrometer Varian Cary 50 Bio	Agilent, Santa Clara, USA
<b><u>Zentrifugation:</u></b>	
24 x 2 ml Rotor, Nr. 7593	Thermo Fisher Scientific, Waltham, USA
6 x 50 ml Rotor, Nr. 3057	Thermo Fisher Scientific, Waltham, USA
Heraeus Biofuge Fresco 17 Tischzentrifuge	Thermo Fisher Scientific, Waltham, USA
Heraeus Biofuge Pico 17 Tischzentrifuge	Thermo Fisher Scientific, Waltham, USA
Heraeus Biofuge Primo R Tischzentrifuge	Thermo Fisher Scientific, Waltham, USA

Heraeus Multifuge 3SR+ Tischzentrifuge	Thermo Fisher Scientific, Waltham, USA
Optima MAX-XP Ultrazentrifuge	Beckman Coulter, Brea, USA
Rotor Fiberlite F10-6x500y	Thermo Fisher Scientific, Waltham, USA
Rotor Fiberlite F14-6x250y	Thermo Fisher Scientific, Waltham, USA
Rotor SS-34	Thermo Fisher Scientific, Waltham, USA
Sorvall RC6 Plus Standzentrifuge	Thermo Fisher Scientific, Waltham, USA

**Sonstige Geräte:**

Eismaschine	Thermo Fisher Scientific, Waltham, USA
Hochdruckhomogenisator M110S	Microfluidics, Newton, USA
JEM-2100 Transmissionselektronenmikroskop	Jeol, Akishima, Japan
ES500W Erlangshen CCD-Kamera	Gatan, Pleasanton, USA
MicroPulser Elektroporationseinheit	Bio-Rad, Hercules, USA
Mikrowelle 7809	Severin, Sundern
Mikrowelle Privileg 30889 KOR-6QB7	IRE Beteiligungs GmbH, Stuttgart
MyCycler Thermocycler	Bio-Rad, Hercules, USA
Rotavac Valve Tec Vakuumpumpe	Heidolph, Schwabach
TProfessional Basic Gradient Thermocycler	Biometra, Jena

**2.1.10 Computersoftware und Online-Tools**

ApE 1.17	M. Wayne Davis, University of Utah, Salt Lake City, USA
BioMath Calculators – $T_M$ (Melting Temperature) Calculations for Oligos	<a href="http://www.promega.com/a/apps/biomath/index.html?calc=tm">http://www.promega.com/a/apps/biomath/index.html?calc=tm</a>
BLAST	<a href="http://blast.ncbi.nlm.nih.gov/Blast.cgi">http://blast.ncbi.nlm.nih.gov/Blast.cgi</a>
Boxshade 3.21	<a href="http://www.ch.embnet.org/software/BOX_form.html">http://www.ch.embnet.org/software/BOX_form.html</a>
Cary WinUV 3.00	Varian (Agilent), Santa Clara, USA
CLC Sequence Viewer 6.5.3	CLC bio, Aarhus, Dänemark
Clustal W	<a href="http://www.ebi.ac.uk/Tools/msa/">http://www.ebi.ac.uk/Tools/msa/</a>
Codon Usage Database	<a href="http://www.kazusa.or.jp/codon/">http://www.kazusa.or.jp/codon/</a>
Creative Suite 5 Design Standard	Adobe, San Jose, USA
DichroWeb	<a href="http://dichroweb.cryst.bbk.ac.uk/html/home.shtml">http://dichroweb.cryst.bbk.ac.uk/html/home.shtml</a>
<i>E. coli</i> Codon Usage Analyzer 2.1	<a href="http://www.faculty.ucr.edu/~mmaduro/codonusage/usage.htm">http://www.faculty.ucr.edu/~mmaduro/codonusage/usage.htm</a>
EndNote X4	Thomson Reuters, New York City, USA
Ettan DIGE Imager 1.0.2.21	GE Healthcare, Chalfont St. Giles, Großbritannien

## 2 Material und Methoden

---

ExPASy Bioinformatics	<a href="http://www.expasy.org/">http://www.expasy.org/</a>
GENTle 1.9.4	Magnus Manske, Universität Köln, Köln
Graphical Codon Usage Analyzer	<a href="http://gcua.schoedl.de/">http://gcua.schoedl.de/</a>
Image Quant 5.2	GE Healthcare, Chalfont St. Giles, Großbritannien
Iris 5.3	Infors HT, Bottmingen, Schweiz
LALIGN-Server	<a href="http://embnet.vital-it.ch/software/LALIGN_form.html">http://embnet.vital-it.ch/software/LALIGN_form.html</a>
mMass 5.1.0	Martin Strohal <sup>[150]</sup>
MPEX <sup>[151]</sup>	<a href="http://blanco.biomol.uci.edu/mpex/">http://blanco.biomol.uci.edu/mpex/</a>
NanoDrop ND-1000 3.7.1	Thermo Fisher Scientific, Waltham, USA
Office Professional 2010	Microsoft, Redmond, USA
Origin 8.1G	OriginLab, Northampton, USA
Protein Database (PDB)	<a href="http://www.rcsb.org/pdb/home/home.do">http://www.rcsb.org/pdb/home/home.do</a>
ProtParam	<a href="http://web.expasy.org/protparam/">http://web.expasy.org/protparam/</a>
PSIPRED	<a href="http://bioinf.cs.ucl.ac.uk/psipred/">http://bioinf.cs.ucl.ac.uk/psipred/</a>
PubMed	<a href="http://www.ncbi.nlm.nih.gov/pubmed">http://www.ncbi.nlm.nih.gov/pubmed</a>
Resource Portal GenBank	<a href="http://www.ncbi.nlm.nih.gov/genbank/">http://www.ncbi.nlm.nih.gov/genbank/</a>
ScanWizard Bio 1.01	Microtek, Hsinchu, Taiwan
Spectra Manager 1.54.03	Jasco, Tokyo, Japan
Spectra Manager 2.08.04	Jasco, Tokyo, Japan
UNICORN 5.11	GE Healthcare, Chalfont St. Giles, Großbritannien
UniProt	<a href="http://www.uniprot.org/">http://www.uniprot.org/</a>
Web of Knowledge 5.8	<a href="http://apps.webofknowledge.com">http://apps.webofknowledge.com</a>

## 2.2 Bioinformatische Methoden

### 2.2.1 Bestimmung der physikalisch-chemischen Eigenschaften von Proteinen

In dieser Arbeit genannte Molekulargewichte, isoelektrische Punkte, molare Extinktionskoeffizienten (für UV-Licht von 280 nm Wellenlänge) und durchschnittliche Hydropathizitätsindices (nach Kyte und Doolittle<sup>[152]</sup>) von Proteinen wurden mittels des Online-Tools ProtParam der ExPASy-Resourcendatenbank ermittelt (<http://web.expasy.org/prot-param/>).

### 2.2.2 Erstellen von DNA- und Aminosäuresequenzalignments

Die Erstellung von Alignments zweier oder mehrerer DNA- bzw. Aminosäuresequenzen ermöglicht einen Vergleich und eine Analyse der Sequenzen hinsichtlich verwandtschaftlicher (phylogenetische/evolutionäre Verwandtschaft und Isoformen auf DNA-Ebene) bzw. struktureller (Homologien und Identifizierung funktioneller Domänen/Bereiche auf Proteinebene) Ähnlichkeiten. Die Erstellung der Alignments erfolgte auf Grundlage der Algorithmen Clustal-W<sup>[153]</sup>, Smith-Waterman<sup>[154]</sup> oder Huang-Miller<sup>[155]</sup>, wobei erster zur Erzeugung globaler und die beiden letztgenannten zur Erzeugung lokaler Sequenzalignments verwendet wurden. Der Clustal-W- und der Smith-Waterman-Algorithmus waren integraler Bestandteil der DNA-Manipulationssoftware GENTle 1.9.4 und der Huang-Miller-Algorithmus war über den LALIGN-Server ([http://embnet.vital-it.ch/software/LALIGN\\_form.html](http://embnet.vital-it.ch/software/LALIGN_form.html)) verfügbar. Zur graphischen Darstellung der Alignments wurde das Online-Tool Boxshade 3.21 ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)) verwendet.

### 2.2.3 Sekundärstrukturvorhersage von Proteinen

Zur Vorhersage von Sekundärstrukturelementen innerhalb der *flank*-Proteine wurde das Online-Tool PSIPRED verwendet. Basierend auf einem neuronalen Netzwerk wird auf Grundlage der Primärstruktur eines Proteins eine Vorhersage über das Auftreten sowohl von unstrukturierten sowie helikalen und  $\beta$ -Faltblattbereichen getroffen. PSIPRED wird online (<http://bioinf.cs.ucl.ac.uk/psipred/>) durch das Bloomsbury Centre for Bioinformatics des University College London zur Verfügung gestellt.

### 2.2.4 Vergleich der Codonverwendung (*codon usage*)

Der genetische Code ist degeneriert, d.h. die meisten Aminosäuren werden durch mehrere verschiedene – sog. synonyme – Basentriplets (Codons) codiert. Im Laufe der Evolution entwickelten sich zwischen verschiedenen Spezies Unterschiede bei der Häufigkeit (d.h. der „Neigung“), mit der das gleiche Codon zur Codierung einer bestimmten Aminosäure verwendet wird<sup>[156,157]</sup>. Eine unterschiedliche Codonneigung zwischen einer

heterolog zu exprimierenden DNA-Sequenz und dem Wirtsorganismus kann (negative) Auswirkungen auf die Effizienz der Expression und somit auch auf die rekombinante Proteinproduktion haben<sup>[157]</sup>. Aus diesem Grund wurde die Codonverwendung in den flankierenden Regionen der *M. galloprovincialis* preCols mit der Codonverwendung des Wirtsorganismus *E. coli* mittels des *Graphical Codon Usage Analyzer*<sup>[158]</sup> (<http://gcua.schoedl.de/>) verglichen. Der Graphical Codon Usage Analyzer bietet die Möglichkeit, die Frequenzen, mit der synonyme Codons innerhalb des Genoms eines Organismus verwendet werden, zu normieren<sup>[159]</sup>. Die infolge der Normierung als „relative Angepasstheit“ bezeichneten Werte erleichtern sowohl die Vergleichbarkeit der Codonverwendung zwischen synonymen Codons verschiedener Gene desselben Organismus wie auch zwischen verschiedenen Organismen.

### 2.3 Molekularbiologische Methoden

#### 2.3.1 Verwendete Lösungen und Puffer

Zur Durchführung der in Kapitel 2.3 beschriebenen Methoden wurden die folgenden Reagenzien, Puffer und Lösungen verwendet. Sofern kein Lösungsmittel angegeben ist, wurde Reinstwasser (im Folgenden als MQ-Wasser bezeichnet) verwendet.

dNTP-Mix	je 10 mM dATP, dCTP, dGTP und dTTP
Enzym-Reaktionspuffer (5 x–10 x)	vom jeweiligen Enzymhersteller bereitgestellt; Zusammensetzung siehe entsprechende Datenblätter
TAE (50 x)	2 M Tris/Essigsäure, pH 8, 50 mM Na-EDTA
DNA-Probenauftragspuffer (6 x)	50 % (v/v) Glycerin, 10 mM EDTA, pH 8, 0,2 % (w/v) Bromphenolblau, 0,2 % (w/v) Xylencyanol

#### 2.3.2 Auftrennung von DNA

Sowohl die analytische wie auch präparative Auftrennung von DNA-Gemischen nach Größe der darin enthaltenen DNA-Moleküle erfolgte mittels Agarose-Gelelektrophorese<sup>[160]</sup>. Je nach benötigter Trennschärfe wurden Agarosegele mit einer



Agarosekonzentration von 0,7 % bis 1,5 % verwendet. Zur Herstellung der Agarosegele wurde die entsprechende Menge Agarose in 1 x TAE-Puffer aufgekocht. Nach Abkühlen der Agaroselösung auf 40–50 °C wurde der nukleinsäurespezifische Farbstoff *SYBR® Safe DNA Gel Stain* in 10 000-facher Verdünnung zugesetzt und die Gellösung in einen horizontalen Gießstand gegossen. In diesem bildete sich durch Abkühlung ein poröses Gel mit einer Porengröße, die von der ursprünglichen Agarosekonzentration abhängig war und im Bereich 50–500 nm lag<sup>[160,161]</sup>. Die zu trennenden DNA-Proben wurden mit entsprechenden Volumina 6 x DNA-Probenauftragspuffer gemischt und in die Taschen des Agarosegels pipettiert. Zur späteren Abschätzung der Größe der DNA-Moleküle wurden zusätzlich ein oder mehrere DNA-Referenzstandards (1 kb DNA-Leiter und/oder 100 bp DNA-Leiter) aufgetragen. Die Auftrennung der DNA-Moleküle erfolgte schließlich in einer horizontalen Agarose-Gelelektrophoresekammer (Serva Electrophoresis) durch Anlegen einer konstanten Spannung (80 V bzw. 120 V, je nach Agarosegelgröße) für 30–60 min. Nach Abschluss der Elektrophorese wurden die DNA-Moleküle im Agarosegel mit dem GelDoc XR Transilluminator visualisiert, indem der mit der DNA wechselwirkende Farbstoff *SYBR® Safe DNA Gel Stain* durch UV-Licht angeregt wurde.

### 2.3.3 Reinigung und Isolierung von DNA

Die Isolierung von Plasmid-DNA und Vektorkonstrukten (d.h. Plasmiden mit integrierten transgenen Nukleotidsequenzen) wurde mit dem *Wizard® Plus SV Minipreps DNA Purification System* (Promega) durchgeführt. Die Durchführung erfolgte gemäß der Standardarbeitsanweisung (SOP, von engl. *standard operating procedure*) des Herstellers. Nur das zur Elution der DNA von den Reinigungssäulchen im letzten Schritt verwendete sterile MQ-Wasser wurde vor Verwendung in der Mikrowelle auf ca. 50 °C vorgewärmt, da auf diese Weise eine Steigerung der Ausbeute erreicht werden konnte.

Zur Reinigung und Isolierung von PCR-Produkten sowie von DNA-Restriktionsverdauen wurde das *Wizard® SV Gel and PCR Clean-Up System* (Promega) nach Hersteller-SOP verwendet. Die Reinigung und Isolierung von DNA wurde entweder direkt von den entsprechenden PCR-Ansätzen ausgehend oder, im Fall von DNA-Gemischen (z. B. nach Restriktionsverdauen), nach agarosegelelektrophoretischer Auftrennung aus einem ausgeschnittenen Agarosegelstück durchgeführt.

### 2.3.4 Polymerasekettenreaktion (PCR)

Die Polymerasekettenreaktion (PCR, von engl. *polymerase chain reaction*) ermöglicht sowohl die Amplifikation als auch die Mutagenese von DNA-Sequenzen *in vitro*<sup>[162]</sup>. Die zu amplifizierende DNA wird thermisch in ihre zwei komplementären Einzelstränge, die dann in folgenden Schritten als Template dienen, denaturiert. Anschließend kommt es zur Anlagerung (engl. *annealing*) kurzer, als Primer bezeichneter DNA-Oligonukleotide, die zum jeweiligen 3'-Ende der beiden Templatstränge komplementär sind. Die Primeranlagerung ist essentiell, da DNA-Polymerasen einen neuen DNA-Strang nicht *de novo* synthetisieren können, sondern auf ein endständiges Nukleotid mit freier 3'-Hydroxylgruppe der Ribose angewiesen sind. Ausgehend von den so entstandenen, partiell doppelsträngigen DNA Bereichen kann eine thermostabile DNA-Polymerase im weiteren Verlauf der PCR-Reaktion nun die Komplettierung der Templateinzelstränge zu doppelsträngiger DNA durchführen. Durch zyklische Wiederholung von Doppelstrangdenaturierung, Primer-Hybridisierung und DNA-Polymerase-vermittelter Strangsynthese wird eine exponentielle Vervielfältigung der ursprünglichen DNA-Sequenz erreicht.

#### 2.3.4.1 Standard-PCR zur Amplifikation von DNA-Sequenzen

Mittels Standard-PCR wurden die für die flankierenden Regionen der preCols codierenden und entweder in der cDNA-Bank bzw. bereits in Plasmiden enthaltenen DNA-Sequenzen amplifiziert.

Basierend auf der Sequenz der zu amplifizierenden DNA wurden Primer entworfen (vgl. Kapitel 2.1.3), welche den 3'-Enden der Templateinzelstränge komplementär waren und an diese hybridisieren konnten. Darüber hinaus wurde jeder Primer an seinem 5'-Ende mit einer zusätzlichen Überhangsequenz ausgestattet. Diese bestand bei den Vorwärtsprimern zum einen aus der Erkennungs- und Schnittsequenz der Restriktionsendonuklease, welche die für die anschließende Ligation benötigten Enden erzeugt (vgl. Kapitel 2.3.5 & 2.3.6). Zum anderen befanden sich 5' der für die enzymatische Restriktion benötigten Nukleotidsequenz zusätzlich noch sechs (beliebige) Basen, da viele Restriktionsnukleasen nicht direkt an das Ende eines DNA-Stranges binden können. Die Überhangsequenzen der Rückwärtsprimer wurden nach dem gleichen Prinzip entworfen,

enthielten jedoch zwischen der für die enzymatische Restriktion benötigten Nukleotidsequenz und dem sechs Basen langen Überhang noch zwei für die Translationstermination codierende Basentriplets, um eine ungewollte Translation von Plasmidsequenzen zuverlässig zu unterbinden. Als DNA-Polymerasen wurden wahlweise die *PfuTurbo*<sup>®</sup> oder die *Phusion*<sup>®</sup> High-Fidelity DNA-Polymerase verwendet, da beide über eine 3' → 5' Exonukleaseaktivität verfügen und Basenfehlpaarungen bereits während der Strangsynthese beheben können. Zur Durchführung der Standard-PCR-Reaktionen wurden folgendes Ansatzschema und folgende Thermocyclerparameter verwendet.

#### Allgemeiner Reaktionsansatz:

Templat-DNA	10–200 ng
Vorwärtsprimer	200 nM
Rückwärtsprimer	200 nM
<i>PfuTurbo</i> <sup>®</sup> - bzw. <i>Phusion</i> <sup>®</sup> High-Fidelity DNA-Polymerase	2,5 U bzw. 2,0 U
Enzym-Reaktionspuffer	1 x
dNTP-Mix	200 µM
MQ-Wasser	ad 100 µl

#### Durchführung:

Anzahl Zyklen:	Dauer [s]:	Temperatur [°C]:	Phase:
1 x	300	95	initialer Denaturierungsschritt
	30	95	Denaturierung
30–35 x	45	$T_M^* - 3\text{ }^{\circ}\text{C}^4$ $T_M^* + 3\text{ }^{\circ}\text{C}^5$	Primer-Anlagerung
	45	72	Elongation
1 x	300	72	verlängerter Elongationsschritt
	nach Bedarf	4	Lagerung

\*Primer-Schmelztemperatur ( $T_M$ ) =  $64,9^{\circ}\text{C} + 41^{\circ}\text{C} \times \frac{(N_G + N_C - 16,4)}{N_S}$  (gemäß dem verwendeten Online-Tool „BioMath Calculators –  $T_M$  (Melting Temperature) Calculations for Oligos“);  $N_G$ : Anzahl der Guanosine;  $N_C$ : Anzahl der Cytidine;  $N_S$ : Anzahl der zwischen Primer und DNA-Sequenz komplementären Basen (ohne Überhänge und Fehlpaarungen). Die zur Durchführung der PCR verwendete Primer-Anlagerungstemperatur entsprach der niedrigeren des Vorwärts-Rückwärts-Primerpaares.

<sup>4</sup> bei Verwendung der *PfuTurbo*<sup>®</sup> DNA-Polymerase

<sup>5</sup> bei Verwendung der *Phusion*<sup>®</sup> High-Fidelity DNA-Polymerase

### 2.3.4.2 Mutagenese-PCR

Die Mutagenese-PCR diente dem gezielten Austausch von Basen während der PCR-Amplifikation einer (zirkulären) DNA-Sequenz. Die verwendeten Primer wurden so entworfen, dass sie die veränderte Nukleotidsequenz sowie 5' und 3' davon zu den jeweiligen Einzelsträngen der DNA-Sequenz komplementäre Sequenzen beinhalteten. Zudem waren beide Primer so entworfen, dass beide an der zu verändernden Stelle – einmal auf dem codierenden und einmal auf dem nichtcodierenden Einzelstrang – binden und es nicht wie bei der Standard-PCR zu einer Begrenzung der zu amplifizierenden Sequenz kommt, sondern das gesamte Plasmid durch die Polymerase amplifiziert werden kann.

In Rahmen dieser Arbeit wurde die Mutagenese-PCR dazu eingesetzt, um aus der eCT-Nukleotidsequenz die *in vivo* vorkommende Erkennungs- und Schnittstelle der Restriktionsendonuklease *Bam*HI (vgl. Kapitel 2.1.6) durch Mutation zu entfernen. Dieser Schritt war nötig, da bei der Erzeugung des Vektorkonstrukts pET28 SUMO-eCT<sub>2Y</sub> (Tab. 2-3) methodenbedingt mit *Bam*HI gearbeitet werden musste. Die Durchführung der Mutagenese-PCR orientierte sich an der SOP des *QuikChange Site-Directed Mutagenesis Kit* der Firma Stratagen.

#### Reaktionsansatz:

Templat-DNA (pET28a-eCT)	20 ng
Primer #1 (eCTBamH1mut2_fwd)	250 nM
Primer #2 (eCTBamH1mut2_rev)	250 nM
<i>PfuTurbo</i> <sup>®</sup> DNA-Polymerase	2,5 U
Enzym-Reaktionspuffer	1 x
dNTP-Mix	200 µM
MQ-Wasser	ad 50 µl

#### Durchführung:

Anzahl Zyklen:	Dauer [s]:	Temperatur [°C]:	Phase:
1 x	120	95	initialer Denaturierungsschritt
	30	95	Denaturierung
20 x	60	55	Primer-Anlagerung
	450	72	Elongation
1 x	600	72	verlängerter Elongationsschritt
	nach Bedarf	4	Lagerung

### 2.3.4.3 Screening-PCR

Per Screening-PCR wurde nach Abschluss von Klonierungen analysiert, ob die auf den entsprechenden Selektionsplatten wachsenden *E. coli*-Zellen nur religierte Plasmide oder aber Plasmide mit inserierter Transgensequenz aufgenommen hatten. Dazu wurde aus den jeweiligen *E. coli*-Zellen die Plasmid-DNA isoliert (vgl. Kapitel 2.3.3) und als DNA-Templat für die Screening-PCRs verwendet. Für Screening-PCRs wurde die *taq* DNA-Polymerase verwendet. Sowohl Vorwärts- als auch Rückwärtsprimer waren in der Regel die gleichen Primer, welche bereits bei der PCR-Amplifikation der entsprechenden Transgensequenzen zum Einsatz kamen. Aus diesem Grund war die Primer-Anlagerungstemperatur auch identisch mit der der Standard-PCR (vgl. Kapitel 2.3.4.1).

#### Allgemeiner Reaktionsansatz:

Templat-DNA	30–60 ng
Vorwärtsprimer	200 nM
Rückwärtsprimer	200 nM
<i>taq</i> DNA-Polymerase	2,5 U
Enzym-Reaktionspuffer	1 x
dNTP-Mix	200 µM
MQ-Wasser	ad 50 µl

#### Durchführung

Anzahl Zyklen:	Dauer [s]:	Temperatur [°C]:	Phase:
1 x	120	94	initialer Denaturierungsschritt
25 x	30	94	Denaturierung
	45	T <sub>M</sub>	Primer-Anlagerung
	45	72	Elongation
	300	72	verlängerter Elongationsschritt
	nach Bedarf	4	Lagerung

### 2.3.5 DNA-Restriktionsverdau

Um DNA-Abschnitte beliebig miteinander kombinieren zu können, müssen diese zunächst mit Restriktionsendonukleasen verdaut werden. Restriktionsendonukleasen erkennen meist 4–8 Nukleotide lange, spezifische Basenabfolgen in doppelsträngiger DNA und hydrolysieren diese an einer bestimmten Stelle. Je nach verwendeter Restriktionsendonuklease entstehen dadurch DNA-Stücke, die an ihren Termini entweder einzelsträngige Überhänge (sog. klebrige Enden) oder aber glatte Enden besitzen. Bei den im Rahmen dieser Arbeit zur Erzeugung von Vektorkonstrukten (vgl. Kapitel 2.1.4) benutzten Restriktionsendonukleasen (vgl. Kapitel 2.1.6) handelte es sich ausschließlich um solche, die nach DNA-Hydrolyse klebrige Enden erzeugen. Besaßen DNA-Fragmente an beiden Enden unterschiedliche klebrige Enden, begünstigte dies die spätere Ligation der entstandenen DNA-Enden in korrekter Orientierung zueinander, da komplementäre klebrige Enden spezifisch miteinander hybridisieren konnten.

Restriktionsverdaue wurden sowohl im analytischen Maßstab zur Identifizierung von Plasmiden mit transgenen Inserts wie auch im präparativen Maßstab zur Erzeugung von DNA-Fragmenten für nachfolgende Ligationen (vgl. Kapitel 2.3.6) durchgeführt.

#### Allgemeiner Reaktionsansatz:

	<u>analytisch:</u>	<u>präparativ:</u>
DNA (Plasmid, Vektorkonstrukt, PCR-Produkt)	300–500 ng	1–5 µg
Restriktionsendonuklease(n)	5–10 U	10–20 U
BSA (optional)	0,1 mg/ml	0,1 mg/ml
Reaktionspuffer	1 x	1 x
MQ-Wasser	ad 20 µl	ad 100 µl

Soweit nach Herstellerangabe möglich, wurden Restriktionsverdaue, für die zwei verschiedene Enzyme verwendet wurden, als Doppelverdau (d. h. mit simultanem Verdau der DNA durch beide Restriktionsendonukleasen) durchgeführt. Verdaut wurde beim jeweiligen Temperaturoptimum der Restriktionsendonukleasen 2–16 h unter leichtem Schütteln (450 U/min) im Tischinkubator. Hitzedeaktivierbare Restriktionsendonukleasen wurden danach 20 min bei 65 °C inaktiviert. War ein Doppelverdau nicht möglich, wurde zuerst 2–3 h mit der Restriktionsendonuklease verdaut, deren Reaktionspuffer

die niedrigeren Salzkonzentrationen besaß. Wann immer möglich, erfolgte direkt im Anschluss die Hitzedeaktivierung der Restriktionsendonuklease. Danach wurde die zweite Restriktionsendonuklease zusammen mit ihrem optimalen Reaktionspuffer zu den Ansätzen hinzugefügt und der Ansatz derart mit MQ-Wasser aufgefüllt, dass der hinzugekommene Reaktionspuffer final erneut einfach konzentriert war. Nach weiteren 2–16 h Inkubation beim Temperaturoptimum der zweiten Restriktionsendonuklease wurde auch diese, falls möglich, thermisch inaktiviert. Alle Restriktionsverdauungen wurden durch Agarose-Gelelektrophorese analysiert (vgl. Kapitel 2.3.2), und für die Erzeugung von Vektorkonstrukten benötigte DNA-Fragmente wurden aus den Agarosegelen isoliert und gereinigt (vgl. Kapitel 2.3.3).

### 2.3.6 Ligation von DNA-Enden

Nach der Hybridisierung komplementärer Enden von DNA-Molekülen können diese durch das Enzym T4 DNA-Ligase wieder kovalent miteinander verbunden werden. Dazu aktiviert das Enzym das freie 5'-Phosphat-Ende des einen DNA-Moleküls durch Adenylierung. Die aktivierte Phosphatgruppe kann anschließend nukleophil durch die freie 3'-Hydroxylgruppe eines weiteren DNA-Moleküls unter Freisetzung des Adenosylmonophosphats angegriffen werden, was zur Ausbildung einer neuen Phosphodiesterbindung zwischen den beiden DNA-Molekülen führt.

Zur Erzeugung der in dieser Arbeit verwendeten Vektorkonstrukte (vgl. Tab. 2-3) wurden mit den entsprechenden Restriktionsendonukleasen verdauete Plasmide sowie das jeweilige auf die gleiche Weise verdauete PCR-Produkt bzw. synthetisierte Oligonukleotid miteinander ligiert.

#### Allgemeiner Reaktionsansatz:

verdauete Plasmid-DNA	50–150 ng
verdauete(s) PCR-Produkt/Insert-DNA	1–10-fach molarer Überschuss (auf die Plasmid-DNA bezogen)
T4 DNA-Ligase	3–6 U (Weiss units)
T4 DNA-Ligase Reaktionspuffer	1 x
MQ-Wasser	ad 20 µl

Nach Zugabe der T4-DNA-Ligase wurden die Reaktionsansätze entweder bei 16 °C für 3 h, bei Raumtemperatur (RT) für 2 h und anschließend bei 4 °C für 16 h oder nur bei 4 °C für 16 h inkubiert. Nach Inkubation wurde der gesamte Ligationsansatz direkt zur Transformation von *E. coli* DH10B Zellen verwendet (vgl. Kapitel 2.4.4).

### 2.3.7 Sequenzierung von DNA

Um zu überprüfen, ob die in das jeweilige Plasmid inserierte Transgensequenz korrekt und in gewünschter Orientierung eingebaut war, wurden die entsprechenden Vektorkonstrukte in Form isolierter Plasmid-DNA (Konzentration 20–100 ng/μl) zur Sequenzierung an die Firma GATC Biotech (Konstanz) geschickt.

Die Qualität der DNA-Sequenzierungen wurde anhand der Chromatogramme mit der Software ApE überprüft und mit dem in der Software GENTle 1.9.4 integrierten Smith-Waterman-Algorithmus (für lokale Alignments) gegen die *in silico* geplante Ursprungssequenz aligned (vgl. Kapitel 2.2.2). Nachdem durch Sequenzierung positive Vektorkonstrukte identifiziert worden waren wurden diese bis zur weiteren Verwendung bei -20 °C gelagert.

## 2.4 Mikrobiologische Methoden

### 2.4.1 Verwendete Lösungen und Puffer

Zur Durchführung der in Kapitel 2.4 beschriebenen mikrobiologischen Methoden wurden die folgenden Medien, Puffer und Lösungen verwendet. Sofern kein Lösungsmittel angegeben ist, wurde MQ-Wasser verwendet.

Ampicillin-Stammlösung (1000 x)	100 mg/ml Ampicillin-Natriumsalz
Chloramphenicol-Stammlösung (1000 x)	50 mg/ml Chloramphenicol (in Ethanol)
Kanamycin-Stammlösung (1000 x)	35 mg/ml Kanamycin-Sulfat
LB Medium (autoklaviert)	1 % (w/v) Trypton, 0,5 % (w/v) Hefeextrakt, 171 mM NaCl, pH 7 (mit NaOH eingestellt), in VE-Wasser gelöst



LB-Agar Medium (autoklaviert)	LB-Medium + 2 % (w/v) Agar-Agar, in VE-Wasser gelöst
Lösung A (sterilfiltriert)	39 mM Natriumacetat/Essigsäure, pH 5,5, 100 mM Calciumchlorid, 70 mM Mangan(II)-chlorid
Lösung A + Glycerin (sterilfiltriert)	15 % (v/v) Glycerin in Lösung A (s.o.)
SOC-Medium (autoklaviert)	2 % (w/v) Trypton, 0,5 % (w/v) Hefeextrakt, 10 mM Natriumchlorid, 2,5 mM Kaliumchlorid, 10 mM Magnesiumchlorid, 20 mM Glucose, 10 mM Magnesiumsulfat (jedoch erst nach dem Autoklavieren zusetzen!)
Tetracyclin-Stammlösung (1000 x)	12,5 mg/ml Tetracyclin (in Ethanol)

#### 2.4.2 Kultivierung und Aufbewahrung von *E. coli*-Zellen

Alle verwendeten *E. coli*-Stämme wurden mit LB-Medium kultiviert. Die Zellen befanden sich dabei entweder direkt im LB-Medium oder wurden auf LB-Agarplatten kultiviert. Den LB-Medien waren, je nach verwendetem *E. coli*-Stamm und eingebrachtem Vektorkonstrukt, zusätzlich noch die dem jeweiligen genomisch codierten bzw. plasmidvermittelten Selektionsmarker entsprechenden Antibiotika in einfacher Konzentration zugesetzt.

Die Kultivierung in Flüssigsuspension erfolgte in Volumen von 4–100 ml des entsprechenden LB-Mediums in Reagenzgläsern bzw. Erlenmeyerkolben. Dabei wurden Medienvolumen bis 50 ml ausgehend von einer *E. coli* Einzelkolonie mit einer sterilen Impföse inokuliert. Anschließend wurden die Zellen 16–20 h unter Schütteln (650 U/min) bei 37 °C inkubiert. Zum Animpfen von Medienvolumen über 50 ml wurde 1/100 des Volumens des anzuimpfenden Mediums aus einer bereits angewachsenen 4 ml-Flüssigkultur überimpft. Die anschließende Inkubationszeit bei 37 °C und 650 U/min hing von der weiteren Verwendung der *E. coli*-Zellen ab und richtete sich nach der Zelldichte der Flüssigkultur.

Zur Kultivierung von *E. coli*-Zellen auf LB-Agarplatten wurde ein Zehntel Volumen bzw. der gesamte Transformationsansatz — je nachdem, ob Vektorkonstrukte bzw. Plasmid-DNA oder aber Ligationsansätze transformiert worden waren — mit einem sterilen Drigalski-Spatel auf dem Nährmedium ausplattiert. Die LB-Agarplatten wurden 12–20 h (abhängig von der zu einem bestimmten Zeitpunkt gegebenen Größe der Einzelkolonien) bei 37 °C inkubiert.

### 2.4.3 Herstellung transformationskompetenter *E. coli*-Zellen

Um *E. coli*-Zellen mit Fremd-DNA (in Form von Plasmiden und Vektorkonstrukten) transformieren zu können, muss die Zellwand der Bakterien für DNA-Moleküle durchgängig gemacht werden. Die Herstellung transformationskompetenter *E. coli*-Zellen wurde nach dem Protokoll des Lehrstuhls Biomaterialien, welches auf dem Verfahren von Sambrook und Russell beruht<sup>[163]</sup>, durchgeführt.

Ausgehend von einer für 16 Stunden bei 37 °C inkubierten 4 ml *E. coli*-Flüssigkultur wurden 100 ml LB-Medium in 1:100-Verdünnung angeimpft. Wurden *E. coli*-Stämme verwendet, die einen genomisch codierten Selektionsmarker besitzen, war sowohl der 4 ml- als auch der 100 ml-Kultur schon das entsprechende Antibiotikum in einfacher Konzentration zugesetzt worden. Die frisch angeimpften 100 ml-Kulturen wurden anschließend bei 37 °C bis zu einer OD<sub>600</sub> von 0,6–0,8 unter Schütteln (150 U/min) inkubiert. Bei Erreichen der entsprechenden Zelldichte wurde mit Hilfe einer sterilen, 1 M MgCl<sub>2</sub>-Lösung eine MgCl<sub>2</sub>-Konzentration von 20 mM eingestellt und die *E. coli*-Zellen weitere 10 min bei 37 °C und 150 U/min inkubiert. Anschließend wurden die *E. coli*-Zellen in sterilen, auf Eis vorgekühlten 50 ml-Zentrifugenröhrchen 5 min bei 4000 x g in einer auf 4 °C vorgekühlten Zentrifuge zentrifugiert. Das Zellpellet wurde in 10 ml (pro 100 ml Ausgangskultur) eiskalter Lösung A resuspendiert und der Ansatz 60 min auf Eis inkubiert. Danach wurden die resuspendierten Zellen erneut 5 min bei 3000 x g und 4 °C zentrifugiert und das Zellpellet in 2 ml (pro 100 ml Ausgangskultur) eiskalter Lösung A + Glycerin aufgenommen. Abschließend wurden die transformationskompetenten Zellen (komplett auf Eis) in 100 µl-Aliquots aufgeteilt, direkt in flüssigem Stickstoff schockgefroren und bis zur Verwendung bei -80 °C gelagert.

#### 2.4.4 Transformation von *E. coli*-Zellen

*In vitro*-Transformationen von *E. coli*-Zellen mit Fremd-DNA wurden entweder gemäß der SOP des Lehrstuhls Biomaterialien oder per Elektroporation durchgeführt.

##### 2.4.4.1 Standardtransformation

Die bei -80 °C gelagerten Aliquots der transformationskompetenten *E. coli*-Zellen wurden zunächst 5 min auf Eis aufgetaut. Danach wurden 20–100 ng Plasmid-DNA bzw. ein Ligationsansatz (20 µl, vgl. Kapitel 2.3.6) zu den angetauten Zellen gegeben, der Transformationsansatz vorsichtig durchmischt und 15 min auf Eis inkubiert. Anschließend wurden die Transformationsansätze für 30–60 s (gemäß Hersteller-SOP für den verwendeten *E. coli*-Stamm) auf 42 °C erwärmt. Nach diesem Hitzeschritt wurden die Ansätze abgekühlt (~ 5 min auf Eis) und pro Ansatz 900 µl LB-Medium (ohne Antibiotikum) hinzugefügt. Die in LB-Medium aufgenommenen Transformationsansätze wurden für 30–60 min bei 37 °C und 750 U/min im Tischinkubator inkubiert. Abschließend wurden die transformierten *E. coli*-Zellen auf LB-Agarplatten mit entsprechendem Antibiotikum ausplattiert. Das Volumen an Zellsuspension, das zum Ausplattieren verwendet wurde, richtete sich dabei nach der zur Transformation verwendeten DNA und betrug 100 µl (nach Transformation von Plasmid-DNA) bzw. das Gesamtvolumen (nach Transformation von Ligationsansätzen). Zum Ausstreichen der gesamten Zellsuspension wurden die Zellen zunächst 5 min bei 5000 x g (RT) sedimentiert, 800 µl des Überstands verworfen und das Zellpellet im verbliebenen Medium resuspendiert und ausplattiert.

##### 2.4.4.2 Elektroporation

Die Transformation von *E. coli*-Zellen mit dem Vektor pET28 SUMO-eCT<sub>2Y</sub> erfolgte, wie in der zur Durchführung der Mutagenese-PCR verwendeten SOP des *QuikChange Site-Directed Mutagenesis Kits* beschrieben (vgl. Kapitel 2.3.4.2), mittels Elektroporation.

Dazu wurden 40 µl transformationskompetenter *E. coli*-Zellen auf Eis aufgetaut und mit 200–500 ng des entsprechenden Vektorkonstrukts vorsichtig in einer auf Eis vorgekühlten Elektroporationsküvette vermischt. Nach einem 4,8 ms langen 1,8 kV-Puls wurden die Elektroporationsansätze sofort in 200 µl vorgekühltem SOC-Medium aufgenommen

und in sterile 1,5 ml Reaktionsgefäße überführt. Anschließend wurden die Transformationsansätze noch 45 min bei 37 °C und 600 U/min im Tischinkubator inkubiert und auf LB-Agarplatten mit dem benötigten Antibiotikum ausplattiert. Die LB-Agarplatten wurden dann 12–16 h bei 37 °C inkubiert.

### 2.5 Rekombinante Produktion von *flank*-Proteinen

#### 2.5.1 Verwendete Lösungen und Puffer

Zur Durchführung der in Kapitel 2.5 beschriebenen Methoden wurden die folgenden Puffer und Lösungen verwendet. Sofern kein Lösungsmittel angegeben ist, wurde MQ-Wasser verwendet.

Antifoam-Stammlösung (1000 x)	1 % (v/v) Breox FMT 30
Fermentationsmedium (autoklaviert)	1,62 % (w/v) Trypton/Pepton aus Casein, 2,48 % (w/v) Hefeextrakt, 15,8 mM Kaliumdihydrogenphosphat, 40,7 mM Ammoniumsulfat, 0,75 mM Calciumchlorid, 4,38 mM Magnesiumsulfat (durch Zugabe von sterilfiltriertem Magnesiumsulfat <u>nach</u> Autoklavieren des Mediums)
IPTG-Stammlösung (1000 x, sterilfiltriert)	1 M IPTG
LB-Medium (autoklaviert)	1 % (w/v) Trypton, 0,5 % (w/v) Hefeextrakt, 171 mM NaCl, pH 7 (mit NaOH eingestellt)
Lysepuffer	20 mM Natriumphosphatpuffer, pH 8, 150 mM NaCl, 0,2 mg/ml Lysozym (direkt vor Gebrauch frisch zugegeben)
Nährlösung (autoklaviert)	80 % (v/v) Glycerin
pH-Korrekturmittel, basisch	2,5 M Natronlauge oder 25 %ige Ammoniaklösung
pH-Korrekturmittel, sauer	2,5 M Phosphorsäure
PMSF-Stammlösung (100 x)	100 mM PMSF in Isopropanol
Waschpuffer	50 mM Tris/HCl, 100 mM Natriumchlorid (pH 7,5)

### 2.5.2 Testexpression

Zur Herstellung der im Rahmen dieser Arbeit verwendeten, rekombinanten *flank*-Proteine wurde zunächst analysiert, welcher *E. coli*-Stamm und welche Bedingungen sich für die heterologe Produktion des jeweiligen Proteins als am geeignetsten erwiesen. Zur Durchführung von Testexpressionen wurden verschiedene *E. coli*-Expressionsstämme (vgl. Kapitel 2.1.5) mit den jeweiligen Vektorkonstrukten transformiert (vgl. Tab. 2-3 & Kapitel 2.4.4), auf LB-Agarplatten mit den benötigten Antibiotika ausplattiert und 12–16 h bei 37 °C inkubiert. Ausgehend von Einzelkolonien wurden 4 ml LB-Flüssigkulturen angelegt und mit dem Verdünnungsfaktor 1:100 auf 100 ml LB-Medien überimpft, wobei den Medien die den Selektionsmarkern entsprechenden Antibiotika bereits zugesetzt waren. Die *E. coli*-Flüssigkulturen wurden bis zum Erreichen einer  $OD_{600} = 0,6–0,8$  bei 37 °C und 150 U/min im Wärmeschrank inkubiert. Anschließend wurden die Flüssigkulturen, falls nötig, auf die jeweilige Expressionstemperatur abgekühlt. Vor sowie an definierten Zeitpunkten nach Induktion der Transgenexpression durch 0,1–1 mM Isopropyl- $\beta$ -D-thiogalactopyranosid (IPTG) wurden 1 ml-Proben aus den *E. coli*-Flüssigkulturen entnommen und deren  $OD_{600}$  bestimmt. Anschließend wurden je 500  $\mu$ l in 1,5 ml-Reaktionsgefäße überführt, die *E. coli*-Zellen 5 min bei 6000 x g und RT zentrifugiert, der Überstand verworfen, das Zellsediment mit zweifach SDS-Auftragspuffer versetzt und 10 min bei 95 °C denaturiert. Das Volumen an verwendetem SDS-Auftragspuffer richtete sich dabei nach der zuvor bestimmten  $OD_{600}$  und betrug 200  $\mu$ l pro  $OD_{600} = 1$ . Nach Beendigung einer Testexpression wurde die zum Endzeitpunkt entnommene 1 ml-Probe zur Bestimmung der Löslichkeit des produzierten rekombinanten Proteins in ihre lösliche und unlösliche Fraktion getrennt. Das Zellsediment wurde dazu statt in SDS-Auftragspuffer in 500  $\mu$ l Lysepuffer resuspendiert. Die resuspendierten *E. coli*-Zellen wurden 60 min bei 4 °C im Lysepuffer inkubiert und danach insgesamt fünfmal bei -20 °C eingefroren und bei RT wieder aufgetaut. Nach abschließender Zentrifugation der Proben (20 000 x g, 30 min, 4 °C) wurde der Überstand mit 125  $\mu$ l fünffach SDS-Auftragspuffer und das Sediment mit 625  $\mu$ l zweifach SDS-Auftragspuffer versetzt. Anschließend wurden alle Proben 10 min bei 95 °C denaturiert. Der Nachweis heterolog gebildeter *flank*-Proteine erfolgte dann mittels SDS-PAGE und Westernblotanalyse (vgl. Kapitel 2.7.2 & 2.7.4).

### 2.5.3 Fermentation transformierter *E. coli*-Stämme

Nachdem durch Testexpressionen ein geeigneter *E. coli*-Expressionsstamm, eine geeignete Expressionstemperatur und eine geeignete Konzentration des Induktionsmittels (IPTG) für die rekombinante Produktion eines bestimmten *flank*-Proteins ermittelt wurden, wurden die zur Proteinproduktion benötigten, transgenen *E. coli*-Zellen durch Batch- bzw. Fed-Batch-Verfahren in größerem Maßstab kultiviert.

#### 2.5.3.1 Kultivierung transgener *E. coli*-Zellen im Batch-Verfahren

Die Kultivierung transgener *E. coli*-Zellen im Batch-Verfahren erfolgte in einem mit Druckluft begasbaren 5 l-Glasgefäß. Sauerstoffgehalt und pH-Wert waren bei diesem Fermentationsverfahren weder unmittelbar messbar noch automatisch regulierbar.

Das 5 l-Glasgefäß wurde mit 4 l LB-Medium gefüllt und dieses direkt zusammen mit dem Bioreaktor autoklaviert. Nach Abkühlen wurde das Medium im Bioreaktor durch ein Wasserbad auf 30 °C temperiert und unter möglichst sterilen Bedingungen je 4 ml der dem Selektionsmarker entsprechenden Antibiotikum-Stammlösung sowie Breox FMT30-Stammlösung zugegeben. Die Inokulation des Bioreaktors erfolgte mit 40 ml einer frischen Übernachtskultur (vgl. Kapitel 2.4.2) des jeweiligen transformierten *E. coli*-Stamms. Anschließend wurden die *E. coli*-Zellen ~ 12 h bei 30 °C inkubiert. Sauerstoff wurde in Form von Druckluft mit einer Flussrate von 3 l/min zugeführt. Die Induktion der Transgenexpression erfolgte durch die während der jeweiligen Testexpression ermittelte IPTG-Konzentration, die durch Zugabe entsprechender Volumina der IPTG-Stammlösung eingestellt wurde. Direkt nach IPTG-Zugabe wurde die Druckluft-Flussrate auf 5–10 l/min erhöht. 4 h nach erfolgter Induktion wurden die *E. coli*-Zellen geerntet (vgl. Kapitel 2.5.4).

#### 2.5.3.1 Kultivierung transgener *E. coli*-Zellen im Fed-Batch-Verfahren

Kontinuierliche Fed-Batch-Verfahren erlauben im Gegensatz zu Batch-Verfahren die Messung und Regulierung des Sauerstoffpartialdruckes ( $pO_2$ ) und des pH-Werts sowie eine kontinuierliche Zuführung von Nährstoffen. Die während dieser Arbeit durchgeführten Fed-Batch-Fermentationen erfolgten alle in 1,3 l Minifors-Bioreaktoren (Infors HT).

Zur Kultivierung der transgenen *E. coli*-Zellen wurde der Bioreaktor mit 1,3 l Fermentationsmedium gefüllt und entsprechend der SOP des Lehrstuhls Biomaterialien vorbereitet. Vor Inokulation des Fermentationsmediums wurde diesem noch das dem jeweiligen Selektionsmarker entsprechende Antibiotikum zugesetzt sowie etwas Nährlösung vorgelegt. Die Inokulation erfolgte durch Zugabe von 40 ml *E. coli*-Zellsuspension aus einer frischen Übernachtskultur (vgl. Kapitel 2.4.2). Die Fermentation der *E. coli*-Zellen verlief semiautomatisiert und wurde mit der Software Iris (Infors HT) überwacht und gesteuert. Temperatur, pO<sub>2</sub>, pH-Wert und Schaumbildung wurden durch im Bioreaktor befindliche Sonden überwacht. Bei unterschreiten des voreingestellten pO<sub>2</sub>-Schwellenwerts von 40 % wurde die Zufuhr von Sauerstoff erhöht und Abweichungen des pH-Sollwerts von 6,8 wurden durch automatische Zugabe von saurem bzw. basischem pH-Korrekturmittel ausgeglichen. Schaumbildung wurde durch automatische Einleitung der Antifoam-Stammlösung reduziert. Die Zugabe der Nährlösung erfolgte manuell und orientierte sich am Sauerstoffverbrauch der *E. coli*-Zellen. Während der Fermentation wurden in regelmäßigen Abständen Proben entnommen und das Zellwachstum über die Bestimmung der Zelldichte überwacht. Bei Erreichen einer Zelldichte, die einer OD<sub>600</sub> von 20–30 entsprach, wurde die Temperatur auf die gewünschte Expressionstemperatur gestellt und die Transgenexpression durch Zugabe von IPTG induziert. Die Ernte der Zellen (vgl. Kapitel 2.5.4) erfolgte in der Regel 4 h nach Induktion.

#### **2.5.4 Ernten und Waschen von *E. coli*-Zellen**

Nach Beendigung der Fermentation wurden die *E. coli*-Zellen geerntet und das Zellpellet zur Entfernung von Medienrückständen gewaschen. Zunächst wurden die Zellen für 30 min bei 8000 x g und 4 °C sedimentiert. Das Zellpellet wurde anschließend mit Hilfe des Laborrührers RZR 2020 (Heidolph) gründlich in 200–250 ml Waschpuffer resuspendiert und erneut zentrifugiert. Je nach Masse des Zellpellets wurden insgesamt 1–3 Waschschriffe durchgeführt. Abschließend wurden die Zellen erneut in Waschpuffer (5 ml pro Gramm Zellpellet) resuspendiert, in flüssigem Stickstoff schockgefroren und bis zur weiteren Verwendung bei -80 °C gelagert.

### 2.5.5 Aufschluss von *E. coli*-Zellen

Die Isolierung und Reinigung von rekombinant in *E. coli* produzierten Proteinen setzt eine Zerstörung der bakteriellen Zellwand voraus. Der Aufschluss der Zellwand von geernteten und gewaschenen *E. coli*-Zellen erfolgte entweder durch Ultraschallbehandlung oder durch Hochdruckhomogenisation.

Dazu wurden bei -80 °C gelagerte, resuspendierte *E. coli*-Zellen zunächst zügig in einem lauwarmen Wasserbad aufgetaut. Nach vollständigem Auftauen wurden pro Milliliter der Zellsuspension ~ 0,2 mg Lysozym (eine entsprechende Menge Lysozym war zuvor in Waschpuffer gelöst worden) sowie, zur Inaktivierung von Proteasen, 1 mM PMSF und 0,3-facher bis max. einfacher Proteaseinhibitor-Mix HP (Serva) zugesetzt und die Zellsuspension 45–60 min auf Eis inkubiert. Zum vollständigen Aufschluss der Zellen durch Ultraschall wurden, je nach Volumen der Zellsuspension, die Ultraschallsonden KE 76 oder TT 13 (Bandelin electronic) verwendet. Die Beschallung erfolgte mit 72 % (KE 76) bzw. 90 % (TT 13) der maximalen Amplitude in sechs Zyklen von je 60 s, wobei nach einem 2 s Puls eine 0,2 s lange sowie zwischen den Zyklen eine 30 s lange Pause erfolgte. Die Zellsuspension wurde während des gesamten Ultraschallaufschlusses kontinuierlich in einem Eiswasserbad gekühlt. Alternativ zum Ultraschallaufschluss erfolgte der Aufschluss der *E. coli*-Zellen auch mechanisch mit hohem Druck im Hochdruckhomogenisator M110S (Microfluidics), wobei zwei Durchläufe mit ~ 0,62 MPa ( $\triangleq$  90 psi) durchgeführt wurden. Unabhängig von der verwendeten Aufschlussmethode wurde das Zelllysate abschließend durch Zentrifugation für 30 min bei 27 000 x g und 4 °C in lösliche und unlösliche Bestandteile getrennt.

### 2.6 Reinigung rekombinant hergestellter Proteine

Die Reinigung eines rekombinant hergestellten Proteins erfolgt durch Abtrennung der im selben Ansatz befindlichen bakteriellen Proteine. Dazu werden spezifische physikalisch-chemische Eigenschaften (Ladung, isoelektrischer Punkt, Hydrophobizität, Löslichkeit, Molekulargewicht, hydrodynamisches Volumen) sowie gegebenenfalls die Affinität des rekombinanten Proteins zu anderen Molekülen ausgenutzt. Die Reinigung der in dieser



Arbeit rekombinant in *E. coli* produzierten *flank*-Proteine erfolgte über verschiedene säulenchromatographische Verfahren.

### 2.6.1 Verwendete Lösungen und Puffer

Zur Durchführung der in Kapitel 2.6 beschriebenen Methoden wurden die folgenden Puffer und Lösungen verwendet. Sofern kein Lösungsmittel angegeben ist, wurde MQ-Wasser verwendet.

CBD-Spaltungspuffer	50 mM Tris/HCl, pH 7,5, 500 mM Natriumchlorid, 50 mM DTT (direkt vor Verwendung frisch zugeben!)
GPC-Laufpuffer	50 mM Tris/HCl, pH 7,5, 100 mM Natriumchlorid
IMAC-A Puffer	40 mM Tris/HCl, pH 7,6, 100 mM Natriumchlorid
IMAC-B Puffer	40 mM Tris/HCl, pH 7,6, 100 mM Natriumchlorid, 500 mM Imidazol
Lyophilisationspuffer	10 mM Ammoniumhydrogencarbonat
Ulp1-Stammlösung	8,3 mg/mL Ulp1, 25 mM Tris/HCl, pH 8, 250 mM NaCl, 500 µM DTT, 50 % (v/v) Glycerin

### 2.6.2 Affinitätschromatographie

Bei der Affinitätschromatographie nutzt man die Affinität des zu reinigenden Proteins zu anderen Molekülen oder funktionellen Gruppen, um dieses von weiteren Proteinen, die nicht die gleiche Affinität zu dem entsprechenden Molekül besitzen, zu trennen. Eine besondere Form der Affinitätschromatographie stellt die ursprünglich von Porath und Kollegen entwickelte Affinitätschromatographie mit immobilisierten Metallionen (IMAC, von engl. *immobilized metal affinity chromatography*) dar<sup>[164]</sup>. Dabei werden Metallionen (v.a.  $\text{Ni}^{2+}$ ) durch einen an einer Festphase immobilisierten Chelator gebunden. Die noch unbesetzten Koordinationsstellen des gebundenen Metallions können dann reversibel durch Moleküle besetzt werden, die als Elektronenpaardonoren (Lewis-Basen) agieren. Bezogen auf Proteine handelt es sich beim Elektronenpaardonor v. a. um die

Stickstoffatome der Seitenkette des Histidins, aber auch Stickstoff- und Schwefelatome in den Seitenketten von Arginin bzw. Cystein können als Lewis-Basen fungieren<sup>[165]</sup>.

Um eine effiziente Bindung der rekombinant erzeugten Proteine an die immobilisierten Metallionen während der IMAC zu gewährleisten, wurden diese als Fusionsproteine mit aminoterminalen SUMO-tag (*s*mall *u*biquitin-like *m*odifier), der eine zusätzliche (His)<sub>6</sub>-Aminosäuresequenz beinhaltet, hergestellt. Als Festphase der IMAC diente die auf Agarosekügelchen basierende Chelating Sepharose Fast Flow-Matrix (GE Healthcare Europe GmbH), an die kovalent der Chelator Iminodiessigsäure (IDA) gekoppelt war. Die durch Komplexierung an IDA-Moleküle immobilisierten Ni<sup>2+</sup>-Kationen besitzen noch drei freie Koordinationsstellen, welche von den freien Elektronenpaaren der Seitenkettenstickstoffatome der Histidinreste der (His)<sub>6</sub>-Aminosäuresequenz besetzt werden können. Nach Bindung des (His)<sub>6</sub>-SUMO-getagten Proteins an die Säulenmatrix und dem Auswaschen nichtbindender (*E. coli*-)Proteine konnte ersteres durch kompetitive Verdrängung, in der Regel durch die Verwendung von Imidazol, wieder von der Säulenmatrix eluiert werden.

Die Reinigung rekombinant hergestellter *flank*-Fusionsproteine wurde an Äkta FPLC Chromatographiesystemen (GE Healthcare) bei ~ 6 °C (im Kühlraum) durchgeführt. Dabei wurden Säulen mit 5–80 ml Bettvolumen verwendet. Zunächst wurde die Chelating Sepharose Fast Flow-Matrix gemäß den Hersteller-SOPs mit Ni<sup>2+</sup> beladen und anschließend mit mindestens zwei Säulenvolumen IMAC-A Puffer äquilibriert, bis die kontinuierlich detektierte UV-Absorption bei einer Wellenlänge von 280 nm ( $A_{280}$ ) eine stabile Grundlinie zeigte. Danach wurde das zu trennende Proteingemisch mit einer Flussrate von 1–5 ml/min (abhängig vom verwendeten Bettvolumen) aufgetragen und die Säule solange weiter mit IMAC-A Puffer gewaschen, bis sich erneut eine stabile  $A_{280}$ -Basislinie einstellte. Um unspezifisch (d. h. mit geringerer Affinität) an die Säulenmatrix gebundene Proteine zu entfernen, folgte ein Waschschriff mit 4–6 % (v/v) IMAC-B Puffer, was einer Imidazolkonzentration von ~ 20–30 mM entsprach. Die abschließende Elution des jeweils zu reinigenden *flank*-Fusionsproteins erfolgte entweder direkt mit 60 % (v/v) IMAC-B Puffer ( $\pm$  300 mM Imidazol) oder, falls bei dieser Variante (wie in Vorversuchen ermittelt) noch eine größere Zahl unspezifisch gebundener Proteine im Eluat vorhanden war, durch einen über 10 Säulenvolumen (SV) reichenden, stufenlosen Gradienten von 0–60 % IMAC-B Puffer ( $\pm$  0–300 mM Imidazol). Eluierte Proteine wurden mit Hilfe eines

Fraktionskollektors in 2–5 ml Fraktionen aufgefangen. Alle Wasch- und Elutionsschritte wurden mit der maximal möglichen Flussrate, bei der der Druck noch unterhalb des Limits von 0,3 MPa lag, durchgeführt. Nach Beendigung der Elution wurden anhand des  $A_{280}$ -Spektrums proteinhaltige Fraktionen ausgewählt und mittels SDS-PAGE und/oder Westernblot (vgl. Kapitel 2.7.2 & 2.7.4) analysiert. Die Reinigung und Regeneration der IMAC-Säulen erfolgte gemäß dem *cleaning-in-place* (CIP)-Protokoll des Herstellers.

### 2.6.3 Gelpermeationschromatographie

Bei der Gelpermeationschromatographie (GPC) beruht die Trennung von Molekülen auf einem Siebeffekt, der durch Interaktionen von Molekülen der mobilen Phase mit der stationären Phase, die aus einer Matrix definierter Porengröße besteht, hervorgerufen wird. Abhängig von ihrem jeweiligen hydrodynamischen Volumen können Moleküle unterschiedlich gut in die Poren der Matrix eindringen. Moleküle mit hydrodynamischen Durchmessern oberhalb der matrixspezifischen Ausschlussgröße können nicht in die Poren eindringen und werden mit der Lösungsmittelfront im Ausschlussvolumen der Säule eluiert. Kleinere Moleküle hingegen dringen in die Poren der Matrix ein und werden, je kleiner ihr hydrodynamischer Durchmesser ist, desto länger von der Matrix zurückgehalten. Die im Proteingemisch enthaltenen Proteine werden folglich entsprechend ihrer hydrodynamischen Durchmesser und somit ihrer Molekülgröße getrennt. Die GPC wird oft als abschließender Schritt bei der Reinigung von Proteinen eingesetzt, um die Reinheit des Zielproteins durch Abtrennung von nicht durch die vorherigen Reinigungsschritte entfernbaren Proteinen sowie von Fragmenten des Zielproteins weiter zu erhöhen<sup>[161]</sup>.

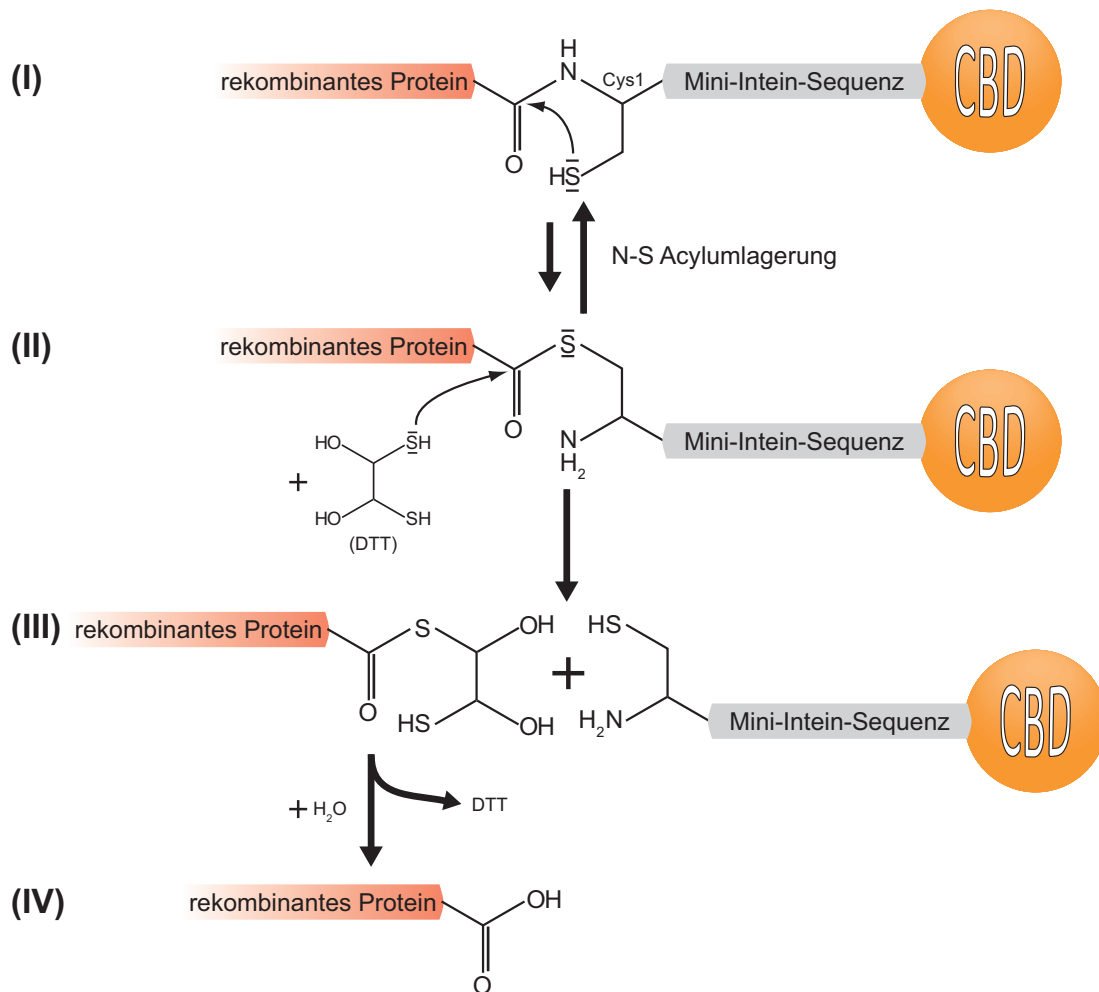
Im Rahmen dieser Arbeit kam die GPC im präparativen Maßstab bei der Reinigung des *flank*-Fusionsproteins cwCT(F34W) zum Einsatz. Die GPC wurde an einem Äkta Purifier Chromatographiesystem (GE Healthcare) durchgeführt und es wurde eine kommerziell erhältliche HiLoad 16/600 Superdex 30 PG Säule (GE Healthcare Europe GmbH) mit einem Säulenvolumen von ~120 ml verwendet. Der optimale Trennbereich der verwendeten Säulenmatrix liegt für Proteine bei 1–10 kDa, so dass Proteine > 10 kDa bereits im Säulenausschlussvolumen eluiert werden. Vor dem Auftragen der Probe wurde die Säule zunächst mit 2 SV MQ-Wasser (Flussrate 1,3 ml/min) gewaschen und

anschließend mit 5 SV GPC-Laufpuffer (Flussrate 1 ml/min) äquilibriert. Anschließend wurde die Probe aufgetragen, wobei das zu applizierende Probenvolumen säulenbedingt max. 4,8 ml ( $\pm 4\%$  des Säulenvolumen) betragen durfte. Um die Anzahl der benötigten GPC-Läufe zu reduzieren, wurden die Proben gegebenenfalls zunächst durch Lyophilisation und erneutem Lösen in kleineren Volumina des GPC-Laufpuffers konzentriert (Proteinendkonzentration  $\sim 3$  mg/ml). Nach dem Auftragen der konzentrierten Probe auf die Gelpermeationssäule erfolgte die Trennung der Proteine bei einer Flussrate von 1 ml/min mit GPC-Laufpuffer. Während der Trennung wurden kontinuierlich die  $A_{280}$  und die UV-Absorption bei 254 nm ( $A_{254}$ ) detektiert. Beim Erreichen eines  $A_{280}$ -Wertes  $> 5$  mAU wurde automatisch mit dem Sammeln von 1,4 ml-Fractionen begonnen. Nach Beendigung der Elution wurden anhand der  $A_{280}$ - und  $A_{254}$ -Spektren proteinhaltige Fraktionen ausgewählt und durch SDS-PAGE und/oder Westernblot (vgl. Kapitel 2.7.2 & 2.7.4) analysiert. Reinigung, Regeneration und Lagerung der Gelpermeationssäule erfolgte gemäß den Herstellerangaben.

### 2.6.4 Abspaltung terminaler Protein-tags

Während der Reinigung rekombinant hergestellter *flank*-Fusionsproteine wurden zwischen den verschiedenen Chromatographieschritten sowohl der aminoterminal SUMO-tag sowie, falls vorhanden, die carboxyterminale Chitin-Bindedomäne (CBD, im Folgenden als CBD-tag bezeichnet), abgespalten, um die entsprechenden tag-freien *flank*-Protein zu erhalten.

Zunächst wurde der carboxyterminale CBD-tag autokatalytisch abgespalten. Die zwischen diesem und der *flank*-Protein Aminosäuresequenz vorhandene bakterielle Mini-Inteinsequenz kann unter Einfluss von Thiolreagenzien autokatalytisch spleißen und dadurch den CBD-tag vom *flank*-Protein abspaltet (Abb. 2-2). Die Abspaltung erfolgt dabei nahtlos und lässt keine vom tag stammenden Aminosäuren am *flank*-Protein zurück<sup>[166,167]</sup>. Zu Beginn wurde die Probe zweimal 12–16 h bei 6 °C gegen je 4 l Intein-Spaltungspuffer dialysiert. Anschließend wurde DTT zur Probe gegeben (DTT-Endkonzentration 50 mM) und diese unter leichtem Rühren für  $\sim 48$  h im Kühlraum inkubiert.



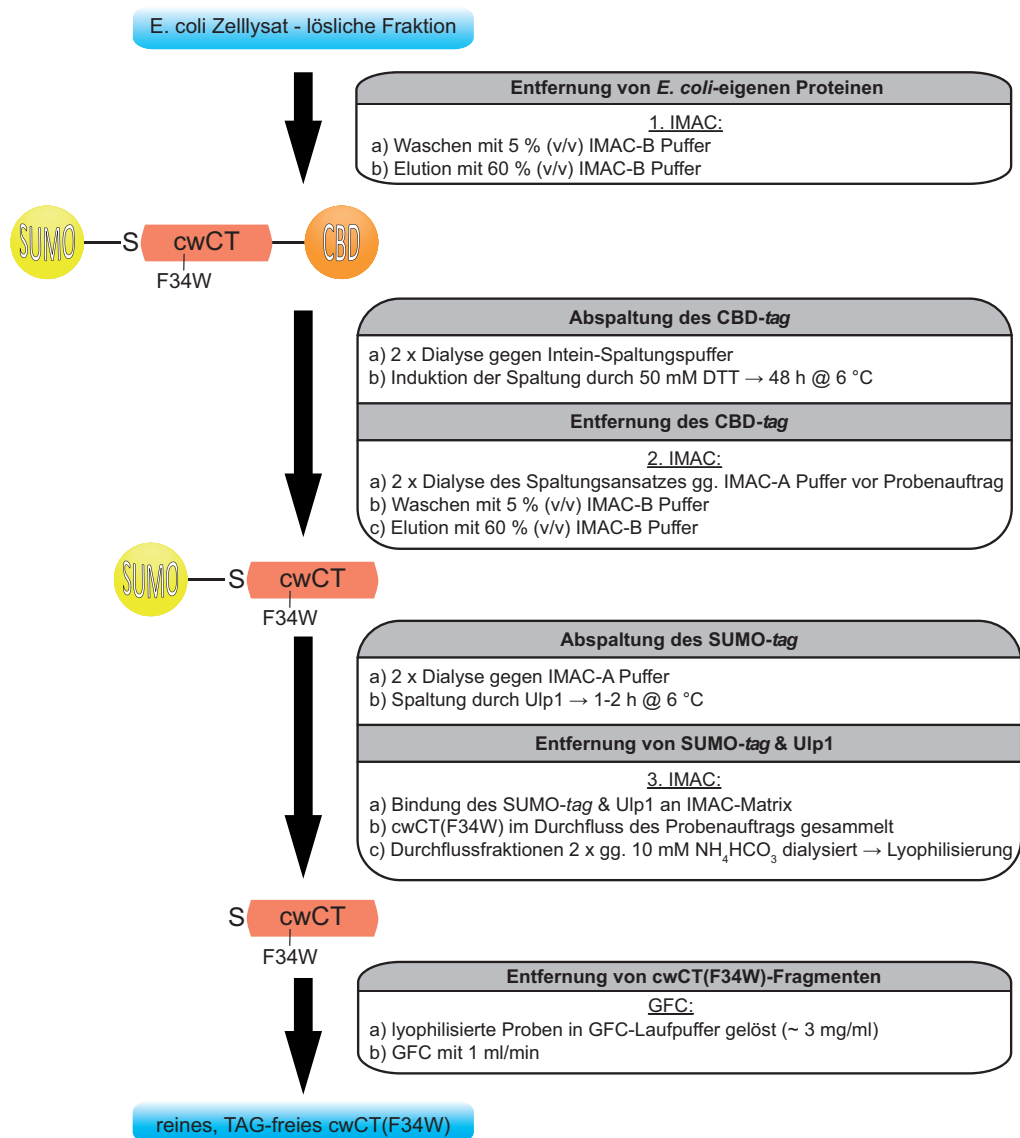
**Abb. 2-2. Reaktionsschema des DTT-vermittelten, autokatalytischen Spleißvorgangs des carboxyterminalen CBD-tags.** (I) Der CBD-tag ist über eine bakterielle Mini-Intein-Sequenz mit dem Carboxyterminus eines rekombinanten *flank*-Proteins verbunden. Durch nukleophilen Angriff des Seitenkettenschwefelatoms des aminoterminalen Cysteinrests der Mini-Intein-Sequenz auf das Carbonylkohlenstoffatom der *upstream* liegenden Aminosäure kann N-S Acylumlagerung auftreten. (II) Dadurch wird aus der ursprünglichen Peptid- eine Thioesterbindung, wobei das Gleichgewicht aber deutlich auf Seite der Peptidbindung liegt. (III) DTT kann jedoch unter Abspaltung der Mini-Intein-Sequenz (und damit auch des CBD-tags) nukleophil an das Carbonylkohlenstoffatom des Thioesters binden und das anfängliche N-Acyl / S-Acyl Gleichgewicht in Richtung des S-Acyls verschieben. (IV) Durch Hydrolyse des *flank*-Protein-DTT-Thioesters entsteht schließlich das tag-freie *flank*-Protein mit nativem Carboxyterminus.

Der abgespaltene CBD-tag konnte, nach Dialyse zur Entfernung des DTT, per erneuter IMAC aus den Ansätzen entfernt werden (vgl. Abb. 2-3 & Kapitel 2.6.2), da er nicht selektiv an die  $\text{Ni}^{2+}$ -Affinitätsmatrix bindet. Bevor die Reinigung fortgesetzt wurde, erfolgte eine Überprüfung der Spaltungseffizienz mittels SDS-PAGE und Westernblot (vgl. Kapitel 2.7.2 & 2.7.4).

Die Abspaltung des SUMO-*tags* erfolgte enzymatisch. Die SUMO-Protease Ulp1 ist in der Lage, spezifisch die Tertiärstruktur von SUMO zu erkennen, ohne dabei auf eine spezifische Erkennungs- oder Schnittsequenz (auf Primärstrukturebene) angewiesen zu sein<sup>[168]</sup>. Diese spezifische Erkennung ermöglicht die effiziente Abspaltung des SUMO-*tags* unter Vermeidung ungewollter proteolytischer Spaltungen, wie sie unter bestimmten Bedingungen bei sequenzabhängigen Protease zu beobachten sind. Ein weiterer Vorteil von Ulp1 ist, dass sie in einem breiten Spektrum von Puffern sowie in Gegenwart von chaotrop oder reduzierend wirkenden Chemikalien aktiv ist<sup>[169]</sup>.

Zur Abspaltung des aminoterminalen SUMO-*tags* der *flank*-Fusionsproteine wurde eine von Lukas Eisoldt produzierte, gereinigte und zur Verfügung gestellte Ulp1-Variante mit aminoterminalen (His)<sub>6</sub>-*tag* verwendet. Nach Abspaltung des SUMO-*tags* konnte Ulp1 auf diesem Wege zusammen mit ersterem per IMAC aus den Spaltungsansätzen entfernt werden. Da mit den entsprechenden Proben vor Abspaltung des SUMO-*tags* schon IMAC-Reinigungsschritte (zur Entfernung der *E. coli*-eigenen Proteine bzw. auch zur Entfernung des abgespaltenen CBD-*tags*) durchgeführt worden waren, wurden die Proben zur Entfernung des enthaltenen Imidazols zunächst 2 x 2 h lang bei ~ 6 °C (im Kühlraum) gegen je 4 l IMAC-A Puffer dialysiert. Anschließend wurde die Spaltung des SUMO-*tags* durch Zugabe der Ulp1-Stammlösung induziert. Die Menge an verwendeter Ulp1-Protease richtete sich dabei nach der zuvor per UV-Spektroskopie bestimmten Gesamteinkonzentration (vgl. Kapitel 2.8.2.2) der Ansätze und es wurde ~ 1 µg Ulp1 pro mg Protein verwendet. Nach Zugabe der Protease wurden die Proben 1–2 h unter leichtem Rühren inkubiert und abschließend Ulp1 sowie der abgespaltene SUMO-*tag* durch IMAC entfernt. Die Überprüfung der Spaltungseffizienz erfolgte durch SDS-PAGE und Westernblot (vgl. Kapitel 2.7.2 & 2.7.4).

## 2.6.5 Reinigungsstrategie für cwCT(F34W)



**Abb. 2-3. Schematische Darstellung der chromatographischen Reinigung von rekombinant produziertem cwCT(F34W).** Ausgehend von der löslichen Fraktion der *E. coli*-Zelllyse wurden die *E. coli*-eigenen Proteine durch IMAC vom Fusionsprotein SUMO-cwCT(F34W)-CBD abgetrennt. Anschließend wurden nacheinander beide tags abgespalten und wiederum durch IMAC vom Protein getrennt. Das tag-freie cwCT(F34W) wurde schließlich durch GPC von cwCT(F34W)-Fragmenten sowie letzten verunreinigenden Proteinen abgetrennt.

Als Prototyp zur Entwicklung einer Reinigungsstrategie rekombinant in *E. coli* produzierter *flank*-Proteine wurde das mit einem aminoterminalen SUMO- und carboxyterminalen CBD-tag versehene *flank*-Fusionsprotein SUMO-cwCT(F34W)-CBD verwendet (vgl. Kapitel 2.1.4). Nach heterologer Expression erfolgte die Reinigung von gebildetem

SUMO-cwCT(F34W)-CBD ausgehend vom löslichen Überstand der aufgeschlossenen *E. coli*-Zellen (vgl. Kapitel 2.5.5). Über die am SUMO-*tag* befindliche (His)<sub>6</sub>-Sequenz wurde das *flank*-Fusionsprotein durch IMAC (vgl. Kapitel 2.6.2) von den *E. coli*-eigenen Proteinen abgetrennt. Anschließend wurden nacheinander der carboxyterminale CBD- und der aminoternale SUMO-*tag* abgespalten (vgl. Kapitel 2.6.4) und ebenfalls (nach Dialysen zur Entfernung des DTT bzw. des Imidazols) per IMAC aus den Ansätzen entfernt. Abschließend wurde zur weiteren Erhöhung der Reinheit des *tag*-freien cwCT(F34W) noch eine GPC durchgeführt (vgl. Kapitel 2.6.3), um verbliebene Proteinverunreinigungen (v. a. durch nicht vollständige Spaltungen im Ansatz verbliebene cwCT(F34W)-Fusionsproteine) von diesem zu trennen (Abb. 2-3).

## 2.7 Proteinanalytische Methoden

### 2.7.1 Verwendete Lösungen und Puffer

Zur Durchführung der in Kapitel 2.7 beschriebenen Methoden wurden die folgenden Puffer und Lösungen verwendet. Sofern kein Lösungsmittel angegeben ist, wurde MQ-Wasser verwendet.

#### SDS-Polyacrylamid-Gelelektrophorese

Anodenpuffer (10 x)	1 M Tris, 0,225 M Salzsäure, pH 8,9
APS-Lösung	10 % (w/v) Ammoniumperoxodisulfat
Gelpuffer-T (3 x) (für <u>T</u> ricin-SDS-PAGE, Verwendung bei der Herstellung der Trenn-T- und Sammelgel-T-Lösungen)	3 M Tris, 1 M Salzsäure, 0,3 % (w/v) SDS, pH 8,45
Kathodenpuffer (10 x)	1 M Tris, 1 M Tricin, pH ~ 8,25 (pH-Wert nicht einstellen, dieser ergibt sich idealerweise aus den enthaltenen Chemikalien), 1 % (w/v) SDS
Sammelgel-G Lösung (für <u>G</u> lycin-SDS-PAGE)	5 % (w/v) Rotiphorese® Gel 40 (37,5 : 1), 125 mM Tris/HCl, pH 6,8, 0,3 % (w/v) SDS
Sammelgel-T-Lösung (für <u>T</u> ricin-SDS-PAGE)	4 % (w/v) Rotiphorese® Gel 40 (29 : 1), 1 M Tris, 0,3 M Salzsäure, 0,1 % (w/v) SDS



SDS-Auftragspuffer (5 x)	312,5 mM Tris/HCl, pH 6,8), 10 % (w/v) Natriumdodecylsulfat (SDS), 50 % (v/v) Glycerin, 0,005 % (w/v) Bromphenolblau, 25 % (v/v) 2-Mercaptoethanol
SDS-Elektrodenpuffer (10 x)	0,25 M Tris, 1,92 M Glycin, 1 % (w/v) SDS
Trenngel-G Lösung (für <u>G</u> lycin-SDS-PAGE)	12,5–15 % (w/v) Rotiphorese® Gel 40 (37,5 : 1), 373 mM Tris/HCl pH 8,8, 0,3 % (w/v) SDS
Trenngel-T-Lösung (für <u>T</u> ricin-SDS-PAGE)	10–16% (w/v) Rotiphorese® Gel 40 (19 : 1), 1 M Tris, 0,3 M Salzsäure, 0,1 % (w/v) SDS

**Silberfärbung**

Entwicklerlösung	2,5 % (w/v) Natriumcarbonat, 0,0148 % (w/v) Formaldehyd
Fixierlösung I	30 % (v/v) Ethanol, 10 % (v/v) Essigsäure
Fixierlösung II	30 % (v/v) Ethanol, 0,5 % (v/v) Essigsäure, 0,4 M Natriumacetat, pH 6, 0,1 % (w/v) Natriumthiosulfat (erst kurz vor Verwendung frisch zugeben!)
Silberfärbelösung	6 mM Silbernitrat, 0,00925 % (w/v) Formaldehyd
Stopplösung	50 mM Natrium-EDTA, pH 7,5

**Westernblot**

Antikörperlösungen	jeweiliger Antikörper 1 : 5000–1 : 10000 in 1 x PBS, optional (laut Hersteller-SOP) mit 0,05 % (v/v) Tween 20 und 1 % Milchpulver, 1 % (w/v) bovinem Serumalbumin (BSA) oder 1 % (w/v) Casein
Blockierlösung	5 % (w/v) Milchpulver in PBS
PBS ( <i>p</i> hosphate <i>b</i> uffered <i>s</i> aline, 10 x)	40 mM Kaliumdihydrogenphosphat, 160 mM di-Natriumhydrogenphosphat, 1,15 M NaCl
PBS-T	0,3 % (v/v) Tween 20 in 1 x PBS
Westernblot-Transferpuffer	25 mM Tris, 150 mM Glycin, 10 % Methanol

**2.7.2 SDS-Polyacrylamid-Gelelektrophorese**

Die SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE) ermöglicht die Trennung von Proteinen nach ihrer Größe (und damit ihrem Molekulargewicht). Die proteinhaltigen Proben werden mit einem denaturierenden Auftragspuffer, welcher Natriumdodecylsulfat

(SDS, von engl. *sodium dodecyl sulfate*) und 2-Mercaptoethanol enthält, versetzt. Das 2-Mercaptoethanol reduziert gegebenenfalls vorhandene Disulfidbrücken und das negativ geladene SDS lagert sich mit einem Verhältnis von 1,4 g SDS pro Gramm Protein an dessen Rückgrat an<sup>[170]</sup>. Die Reduktion der Disulfidbrücken und die Anlagerung der SDS-Moleküle führen zur Denaturierung der in der Probe enthaltenen Proteinmoleküle. Darüber hinaus kommt es durch die Bildung von SDS-Proteinkomplexen zu einer Maskierung der Eigenladung der Proteine und einer gleichmäßigen, der Größe des Proteins in etwa proportionalen negativen Nettoladung. Die dadurch entstandene gleichmäßige Ladung der Proteine ermöglicht es, diese nun elektrophoretisch in einem Polyacrylamid-Gel definierter Porengröße voneinander zu trennen, da die Mobilität nicht mehr von der Eigenladung der Proteine sondern nur noch von deren Molekülgröße abhängt<sup>[170]</sup>. Die elektrophoretische Laufweite eines jeden Proteins innerhalb des Gels verhält sich dabei annähernd proportional zum Logarithmus seines Molekulargewichts.

### 2.7.2.1 Diskontinuierliche Tris-Chlorid/Tris-Glycin SDS-PAGE

Diese auf einem von Davis & Ornstein<sup>[171,172]</sup> entwickelten Elektrophoreseverfahren basierende und von Laemmli<sup>[173]</sup> verfeinerte SDS-PAGE-Methode (im Folgenden als Glycin-SDS-PAGE bezeichnet) verwendet ein diskontinuierliches Tris-HCl / Tris-Glycin Puffersystem sowie zwei unterscheidbare Polyacrylamidgel (PAG)-Bereiche. Im Sammelgel liegt bei einem pH Wert von 6,8 ein Großteil des im Elektrodenpuffer enthaltenen Glycins als Zwitterion vor. Dadurch ist seine elektrophoretische Mobilität gering und es wandert im elektrischen Feld langsamer Richtung Anode als die negativ geladenen SDS-Proteinkomplexe. Letztere migrieren entsprechend ihrer jeweiligen Mobilität zwischen dem zwitterionischen Glycin und den kleinen Chloridanionen. Ein dadurch entstehender Feldstärkegradient führt dazu, dass sich die Proteine übereinander stapeln und der Proteinstapel nun solange isotachophoretisch durch das Sammelgel wandert, bis die Proteine nach und nach den zweiten PAG-Bereich, das Trenngel, erreichen<sup>[161]</sup>. Da das Trenngel deutlich kleinere Poren als das Sammelgel besitzt, kommt es zu einer zusätzlichen Fokussierung der Proteinstapel an der Grenze beider Gele. Der basischere pH-Wert des Trenngels führt dazu, dass sich die elektrophoretische Mobilität des Glycins erhöht. Dadurch existiert innerhalb des Trenngels kein Feldstärkegradient mehr und die Proteine

bewegen sich nun entsprechend ihrer durch Molekülgröße und PAG-Porengröße bestimmten spezifischen Mobilität anisotachophoretisch, was zu ihrer größenabhängigen Trennung führt<sup>[161]</sup>.

Im Rahmen dieser Arbeit wurden kommerzielle 12 %ige und 4–20 %ige Polyacrylamid-Gradientengele der Firma Serva electrophoresis sowie selbst aus den entsprechenden Trenngel-T-Lösungen hergestellte Trenngele (10%, 12,5 % und 15 % Acrylamidkonzentration) mit 5 %igen Sammelgelen verwendet. Die Polymerisation der Trenngele (sowie später auch der Sammelgele) wurde durch Zugabe von 1 % (v/v) APS- und 0,2 % (v/v) TEMED initiiert. Vollständig polymerisierte Trenngele wurden anschließend mit einem ca. zwei Zentimeter breiten, 5 %igen Sammelgel – ausgehend von der Sammelgel-G-Lösung – überschichtet. Die elektrophoretische Auftrennung der zuvor 5–10 min in SDS-Auftragspuffer bei 95 °C denaturierten Proben erfolgte in vertikalen Laukammern in SDS-Elektrodenpuffer. Unabhängig von der Porengröße des verwendeten PAGs wurde zur Durchführung der Elektrophorese 30–60 min eine konstante Stromstärke von 25 mA pro Gel angelegt. Nach Beendigung der Elektrophorese erfolgte eine Analyse durch Silberfärbung und/oder per Westernblot (vgl. Kapitel 2.7.3 & 2.7.4).

### 2.7.2.2 Diskontinuierliche Tris-Chlorid/Tris-Tricin SDS-PAGE

Der Vorteil einer von Schägger & von Jagow entwickelten diskontinuierlichen SDS-PAGE-Methode<sup>[174,175]</sup> (im Folgenden als Tricin-SDS-PAGE bezeichnet) im Vergleich zur Glycin-SDS-PAGE liegt darin, dass Proteine mit Molekulargewichten kleiner 30 kDa besser aufgetrennt werden. Dazu werden bei der Tricin-SDS-PAGE statt eines glycinhaltigen Elektrodenpuffers unterschiedliche Puffersysteme für Kathode (Kathodenpuffer) und Anode (Anodenpuffer) verwendet. Der Kathodenpuffer enthält Tricin, welches beim pH-Wert des hier verwendeten Sammelgels eine höhere elektrophoretische Mobilität als Glycin (unter den pH-Bedingungen der klassischen Glycin-SDS-PAGE) besitzt. Dies führt dazu, dass kleine Proteine sich im Sammelgel nicht wie bei der Glycin-SDS-PAGE innerhalb des Stapels der in Richtung Anode voraus laufenden SDS- und Chloridanionen aufhalten, sondern eigene Stapel mit scharfen Banden bilden und deshalb eine effizientere Vortrennung niedermolekularer Proteine und Peptide stattfindet.

Die für die Tricin-SDS-PAGE benutzten SDS-PAGs bestanden aus einem 16 %igen Trenngel, einer darüber liegenden, ~ 1 cm breiten, 10 %igen Gelzwischenschicht sowie einem 4 %igen Sammelgel. Zunächst wurde das 16 %ige Trenngel aus der Trenngel-T-Lösung gegossen, indem deren radikalische Polymerisation durch Zugabe von 0,4 % (v/v) APS-Lösung und 0,04 % (v/v) TEMED gestartet und diese zwischen zwei in einem vertikalen Gießstand eingespannte Glasplatten pipettiert wurde. Die noch flüssige Gellösung wurde mit MQ-Wasser überschichtet. Nach 30–60 min, frühestens jedoch nachdem sich das 16 %ige Trenngel verfestigt hatte, wurde das MQ-Wasser abgekippt und die 10 %ige Zwischenschicht auf die gleiche Weise gegossen. Die Polymerisation wurde jedoch durch Zugabe von 0,5 % (v/v) APS-Lösung und 0,05 % (v/v) TEMED gestartet. Abschließend wurde das 4 %ige Sammelgel ausgehend von der Sammelgel-T-Lösung unter Zugabe von 1 % (v/v) APS-Lösung und 0,1 % (v/v) TEMED auf der 10 %igen Zwischenschicht polymerisiert. Vor Durchführung der Tricin-SDS-PAGE wurden die mit SDS-Auftragspuffer versetzten Proben für 5–10 min bei 95 °C erhitzt und anschließend (probenabhängig) entsprechende Mengen in die Geltaschen pipettiert. Die Tricin-SDS-PAGE wurde in einer vertikalen Gellaufkammer, die kathodenseitig mit Kathoden- und anodenseitig mit Anodenpuffer gefüllt war sowie über den Kühlwasserkreislauf gekühlt wurde, durchgeführt. Zunächst wurde eine konstante Spannung von 80 V angelegt. Nachdem die Proben in das 10 %ige Zwischengel eingelaufen waren, wurde die Spannung bis zum Ende der Tricin-SDS-PAGE auf 200 V erhöht. Nach Beendigung der Elektrophorese erfolgte die Analyse durch Silberfärbung und/oder Westernblot (vgl. Kapitel 2.7.3 & 2.7.4).

### 2.7.3 Silberfärbung

Bei der Silberfärbung handelt es sich um eine sehr sensitive Färbemethode für Proteine, mit der unter idealen Bedingungen noch Proteinmengen im Pikogrammmaßstab nachweisbar sind<sup>[176,177]</sup>. Die Silberfärbung kam im Rahmen dieser Arbeit standardmäßig zur Färbung und Visualisierung von Proteinen nach SDS-PAGE zum Einsatz.

Nach Beendigung der SDS-PAGE (vgl. Kapitel 2.7.2) wurden die PAGs zunächst kurz mit MQ-Wasser abgespült, um anhaftenden Elektrodenpuffer zu entfernen. Anschließend wurden die Proteine fixiert. Dazu wurden die PAGs zunächst für 10 min in Fixierlösung I und anschließend für mindestens 30 min in Fixierlösung II unter leichtem Schwenken

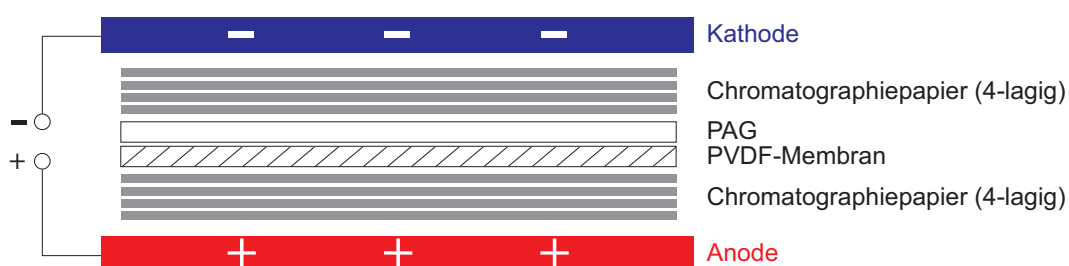
inkubiert. Die Inkubation in Fixierlösung II konnte wahlweise auch über Nacht erfolgen. Optional wurde den beiden Fixierlösungen noch zusätzlich 2 % (v/v) Glutardialdehyd zugesetzt, um die niedermolekularen Proteine zusätzlich quervernetzten und ihre Fixierung im PAG dadurch zu verstärken. Diese zusätzliche Quervernetzung wurde i. d. R. bei Silberfärbungen nach Tris-SDS-PAGE durchgeführt. Nach den Fixierungsschritten wurden die PAGs 4 x 5 min unter leichtem Schwenken in MQ-Wasser gewaschen. Die Färbung der Proteine erfolgte durch 30–45-minütige Inkubation der PAGs in 100 ml Silberfärbelösung (je Gel) unter Lichtausschluss. Das in der Färbelösung enthaltene kationische Silber wird dabei am Proteinrückrat und an funktionellen Gruppen der Aminosäureseitenketten teilweise reduziert, so dass dort Keime aus elementarem Silber entstehen. Durch anschließende Zugabe der reduzierend wirkenden Entwicklerlösung wird weiteres elementares Silber, vorzugsweise an den schon gebildeten Keimen, abgeschieden<sup>[161]</sup> und die im PAG fixierten Proteine auf diese Weise gefärbt. Nach Erreichen der gewünschten Färbungsintensität wurde die Entwicklerlösung abgekippt und die Silberfärbung durch Zugabe von Stopplösung, welche die Reduktion weiterer Silberkationen durch deren Komplexbildung unterbindet, beendet. Die gefärbten PAGs wurden abschließend zu Dokumentationszwecken am Gelscanner HP Scanjet G4050 (Hewlett Packard) eingescannt.

#### **2.7.4 Westernblot**

Beim Westernblot handelt es sich um eine immunochemische Methode, mit der die Detektion bestimmter Proteine nach SDS-PAGE möglich ist. Die Detektion erfolgt dabei durch die spezifische Bindung eines gegen ein Epitop des zu detektierenden Proteins gerichteten Antikörpers. Gebundene Antikörper können anschließend entweder direkt über ein daran konjugiertes Enzym oder Fluoreszenzfarbstoffe oder indirekt über die Bindung eines gegen den primären Antikörper gerichteten und mit einem Enzym oder Fluoreszenzfarbstoff konjugierten sekundären Antikörpers visualisiert werden.

Zunächst wurden die durch SDS-PAGE getrennten Proteine elektrophoretisch auf eine Polyvinylidenfluorid (PVDF)-Membran übertragen. Zur Aktivierung wurde die auf Größe des PAGs zurechtgeschnittene PVDF-Membran für ~ 15 s in 100 %iges Methanol eingetaucht und anschließend für 2 min unter leichtem Schwenken in MQ-Wasser inkubiert.

Die aktivierte PVDF-Membran wurde zusammen mit dem PAG für 15 min in Westerblot-Transferpuffer inkubiert. Kurz vor Ende der Inkubationsperiode wurden acht ebenfalls auf die Größe des PAGs zurechtgeschnittene 3MM CHR Chromatographiepapiere mit Westerblot-Transferpuffer getränkt. Anschließend wurde aus den Chromatographiepapieren, der PVDF-Membran sowie dem PAG ein Stapel mit der in Abb. 2-4 gezeigten Anordnung gebildet und der Stapel luftblasenfrei zwischen den beiden Elektroden der Fastblot-Apparatur (Biometra) platziert.



**Abb. 2-4. Aufbau des Stapels während des elektrophoretischen Transfers von Proteinen aus SDS-Polyacrylamidgelen auf PVDF-Membranen.** Das PAG sowie die PVDF-Membran befinden sich zwischen je 4 Lagen von in Westernblot-Transferpuffer getränkten 3MM CHR Chromatographiepapieren. Die PVDF-Membran befindet sich anodenseitig des PAG.

Durch Anlegen einer konstanten Stromstärke von  $1,25 \text{ mA/cm}^2$  für 60–90 min wurden die Proteine elektrophoretisch vom PAG auf die PVDF-Membran übertragen. Nach Beendigung des Transfers wurde die PVDF-Membran 1–16 h in Blockierlösung geschwenkt. Anschließend wurde die PVDF-Membran mit dem gemäß der Hersteller-SOP verdünnten primären Antikörper 45–90 min unter Schwenken (und im Falle fluoreszenzmarkierter Antikörper unter Lichtausschluss) inkubiert und danach 4 x 5 min mit PBS-T gewaschen. War der primäre Antikörper direkt mit einer Peroxidase oder einem Fluoreszenzfarbstoff konjugiert, wurde die PVDF-Membran 5 min mit PBS gewaschen und der Antikörper anschließend detektiert. Bei Verwendung peroxidasekonjugierter Antikörper wurde die PVDF-Membran gemäß den Herstellerangaben mit dem *Amersham ECL Plus Western Blotting Detection System* (GE Healthcare) behandelt. Die Detektion des dabei von der Peroxidase erzeugten fluoreszierenden Zwischenprodukts erfolgte mit dem *Deep Purple 2*-Filtersatz des *Ettan DIGE* Fluoreszenzbildgebers (GE Healthcare) bei einer Anregungswellenlänge von 390 nm (Spaltbreite 20 nm) und einer Emissionswellenlänge von

595 nm (Spaltbreite 25 nm). Bei Verwendung fluoreszenzmarkierter Antikörper wurde anstelle des *Deep Purple 2*- der Cy2-Filtersatz mit einer Anregungswellenlänge von 480 nm und einer Emissionswellenlänge von 530 nm (Spaltbreite 40 nm) verwendet.

War der primäre Antikörper nicht mit einer Peroxidase oder einem Fluoreszenzfarbstoff konjugiert, wurde die PVDF-Membran nach dem Waschen mit PBS-T mit dem gegen den primären Antikörper gerichteten und ebenfalls nach Hersteller-SOP verdünnten sekundären Antikörper 45 min unter leichtem Schwenken inkubiert und erneut 4 x 5 min mit PBS-T sowie weitere 5 min mit PBS gewaschen. Die abschließende Detektion erfolgte analog dem Vorgehen zur Detektion direkt konjugierter Primärantikörper.

## 2.8 Spektroskopische Methoden

### 2.8.1 Verwendete Lösungen und Puffer

Zur Lösung der Bestandteile des bei einigen der in Kapitel 2.8 beschriebenen Methoden verwendeten künstliche Meerwassers<sup>[108]</sup> (s.u.) wurde MQ-Wasser verwendet.

künstliches Meerwasser (KMW)

450 mM Natriumchlorid, 28 mM Magnesiumsulfat-Heptahydrat, 24 mM Magnesiumchlorid-Hexahydrat, 16 mM Calciumchlorid-Hexahydrat, 2,4 mM Natriumhydrogencarbonat

### 2.8.2 UV-Absorptionsspektroskopie

Abhängig von ihrer Elektronenkonfiguration sind Moleküle in der Lage, elektromagnetische Strahlung bestimmter Wellenlänge(n) zu absorbieren, was bedeutet, dass die Lichtintensität von Licht der entsprechenden Wellenlänge nach Durchgang durch die Probe geringer ist als die des ursprünglich eingestrahltten Lichts. Neben der Elektronenkonfiguration des absorbierenden Moleküls haben noch dessen Konzentration sowie die Weglänge, welche das Licht durch die Probe hindurch zurücklegen muss, einen Einfluss auf das Ausmaß der auftretenden Absorption.

Dieser Zusammenhang wird durch das Lambert-Beer'sche Gesetz beschrieben<sup>[178]</sup>:

**Formel 2-1**       $A(\lambda) = \log\left(\frac{I_0}{I}\right) = \varepsilon(\lambda) \times c \times d$

$A(\lambda)$ : Absorption von Licht der Wellenlänge  $\lambda$   
 $I_0$ : Intensität des eingestrahnten Lichts  
 $I$ : Intensität des ausfallenden Lichts  
 $\varepsilon(\lambda)$ : molarer Absorptionskoeffizient [ $\text{l} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ ] (für Licht der Wellenlänge  $\lambda$ , Moleküleigenschaft)  
 $c$ : Stoffmengenkonzentration [ $\text{mol/l}$ ]  
 $d$ : Schichtdicke [ $\text{cm}$ ] ( $\triangleq$  Weglänge des Lichtes durch die Probe)

Bei der UV-Spektroskopie wird die Absorption elektromagnetischer Strahlung im UV-Spektralbereich (380–100 nm) analysiert. Bei konstanter Schichtdicke  $d$  verhält sich die UV-Absorption eines bestimmten Moleküls also direkt proportional zu seiner Konzentration. Dies hat zur Folge, dass bei bekanntem molarem Absorptionskoeffizient  $\varepsilon$  Konzentrationsbestimmungen durch UV-Absorptionsspektroskopie vorgenommen werden können. Wird die UV-Absorption über einen definierten Spektralbereich aufgenommen, können die auf diese Weise erhaltenen Absorptionsspektren zudem Aufschlüsse über die Identität eines Analyten sowie dessen Reinheit geben.

### 2.8.2.1 UV-Absorptionsspektroskopie an DNA

Zur Konzentrationsbestimmung von DNA in Lösung (Plasmide, Vektorkonstrukte, PCR-Produkte) wurden 2–3  $\mu\text{l}$  der zu untersuchenden Probe verwendet und mit dem NanoDrop ND-100 Photometer (Thermo Fisher Scientific) UV-Absorptionsspektren im Spektralbereich 230–350 nm aufgezeichnet. Der Absorptionswert bei 260 nm ( $A_{260}$ ) wurde von der mit dem Photometer verbundenen Software NanoDrop ND-1000 3.7.1 (Thermo Scientific) direkt zur Bestimmung der DNA-Konzentration verwendet, wobei ein molarer Absorptionskoeffizient  $\varepsilon$  von  $0,02 \text{ (mg/l)}^{-1} \text{cm}^{-1}$  für doppelsträngige DNA verwendet wurde. Da doppelsträngige DNA und Proteine UV-Absorptionsmaxima bei unterschiedlichen Wellenlängen (260 nm gegenüber 280 nm) aufweisen, war durch Bildung des Quotienten  $A_{260}/A_{280}$  eine Überprüfung der Reinheit isolierter DNA (vgl. Kapitel 2.3.3) möglich. Als Richtwert gilt dabei, dass nicht mit Proteinen verunreinigte DNA einen  $A_{260}/A_{280}$ -Quotient  $\geq 1,8$  besitzt<sup>[179,180]</sup>.



### 2.8.2.2 UV-Absorptionsspektroskopie an Proteinen

In Proteinen sind die Seitenketten der aromatischen Aminosäuren Tryptophan, Tyrosin und Phenylalanin sowie eventuell vorhandene Disulfidbrücken aufgrund ihrer Elektronenkonfiguration in der Lage, elektromagnetische Strahlung im UV-Spektralbereich zu absorbieren. Den Hauptbeitrag liefern dabei die Seitenketten der Tryptophane und Tyrosine sowie die Disulfidbrücken, während die UV-Absorption der Phenylalaninseitenketten vernachlässigbar ist.

Die Bestimmung der Proteinkonzentration erfolgte mit dem NanoDrop ND-100 Photometer (Thermo Fisher Scientific). 2–3 µl der zu analysierenden Probe wurden zur Aufnahme eines UV-Absorptionsspektrums im Spektralbereich von 230–350 nm verwendet. Ausgehend vom  $A_{280}$ -Wert konnte anschließend gemäß dem Lambert-Beer'schen Gesetz die Proteinkonzentration der Probe berechnet werden. Der für die Bestimmung verwendete molare Absorptionskoeffizient  $\epsilon$  war dabei von der Aminosäurezusammensetzung der untersuchten Proteine abhängig und wurde basierend auf den individuellen Beiträgen der UV-Licht-absorbierenden Komponenten durch das Online-Tool ProtParam ([www.expasy.org](http://www.expasy.org)) berechnet<sup>[181,182]</sup>. Während der Reinigung rekombinant hergestellter *flank*-Proteine konnte durch Betrachtung des  $A_{260}/A_{280}$ -Quotienten zudem überprüft werden, ob noch DNA-Verunreinigungen in den Proteinproben vorhanden waren<sup>[180]</sup>.

### 2.8.3 Zirkulardichroismusspektroskopie

Die Zirkulardichroismusspektroskopie (CD-Spektroskopie) ermöglicht die konformationelle Analyse von Proteinen hinsichtlich ihrer Sekundär- und Tertiärstruktur. Wie alle chiralen Moleküle zeigen auch aus L-Aminosäuren aufgebaute Proteine eine unterschiedlich starke Absorption von rechts- und linkszirkular polarisiertem Licht. Im Fern-UV-Spektralbereich tragen v. a. die Peptidbindungen des Proteinerückgrats durch Wechselwirkung mit der elektrischen Komponente des zirkularpolarisierten UV-Lichts zu dessen Absorption bei und die Fern-UV-CD-Spektren werden von den Banden der  $n \rightarrow \pi^*$  und  $\pi \rightarrow \pi^*$  Übergänge dominiert<sup>[178,183]</sup>. Je nach Konformation des Proteinerückgrates besitzen die Übergangsdipolmomente der Elektronenübergänge unterschiedliche

räumliche Orientierungen, was zu konformationsspezifischen Spektren führt<sup>[178,183]</sup>. Auf diese Weise lassen sich aus dem Fern-UV-CD-Spektrum eines Proteins Informationen über das Vorkommen von  $\alpha$ -Helices,  $\beta$ -Faltblättern,  $\beta$ -Schleifen sowie ungeordneter *random coil*-Bereiche ziehen. Fern-UV-CD-Spektren klassischer *random coil*-Proteine zeigen dabei ein Minimum bei 195 nm und ein schwaches Maximum bei  $\sim 220$  nm.  $\alpha$ -helikale Proteine weisen ein Maximum bei 191–193 nm sowie Minima bei 208 nm und 222 nm auf, während  $\beta$ -faltblattreiche Proteine ein Maximum bei 195–200 nm sowie ein Minimum bei 218 nm zeigen<sup>[184,185]</sup>. Durch den Vergleich eines beliebigen Fern-UV-CD-Spektrums mit solchen in Datenbanken hinterlegten Prototyp-CD-Spektren, die von Proteinen mit bereits aufgeklärter Struktur detektiert wurden, ist es möglich, sowohl qualitative als auch quantitative Rückschlüsse auf Sekundärstrukturelemente innerhalb des unbekannten Proteins zu ziehen.

Im Rahmen dieser Arbeit wurde die Fern-UV-CD-Spektroskopie zur Analyse der Sekundärstruktur des rekombinant in *E. coli* hergestellten *flank*-Fusionsproteins cwCT(F34W) in verschiedenen wässrigen Lösungsmitteln sowie zur Überprüfung struktureller Auswirkungen der Interaktion desselben mit SDS-Mizellen oder kleinen unilamellaren Lipidvesikeln (SUVs, von engl. *small unilamellar vesicles*) verwendet. Zur Analyse der Sekundärstruktur wurde gereinigtes, lyophilisiertes cwCT(F34W) in Konzentration von  $\sim 15 \mu\text{M}$  in MQ-Wasser, künstlichem Meerwasser (KMW) oder 5 mM MOPS, pH 7,5, bei RT gelöst. Nach Zugabe der entsprechenden Volumina des jeweiligen Lösungsmittels wurden die Proben 1 h vorsichtig in einem Überkopfschüttler gemischt und danach bei 4 °C 45 min lang in einer OPTIMA MAX XP Ultrazentrifuge (Beckman Coulter) bei 186 000 x g zentrifugiert, um unlösliche Bestandteile zu entfernen. Nach Zentrifugation wurde die cwCT(F34W) Konzentration im Überstand UV-spektroskopisch bestimmt (vgl. Kapitel 2.8.2.2) und anschließend Fern-UV-CD-Spektren (von 260 nm bis zum probenbedingten Detektorlimit von  $\sim 195$ –200 nm) der Proben an einem Jasco J-815 Spektropolarimeter (Jasco) gemessen. Datenpunkte wurden mit einer Schrittweite von 0,2 nm und einer Scan-Geschwindigkeit von 50 nm/min bei 2 s Integrationszeit aufgenommen, wobei zur Verbesserung des Signal-Rausch-Verhältnisses jeweils 5 Spektren akkumuliert und gemittelt wurden. Zur Untersuchung der Temperaturabhängigkeit der cwCT(F34W)-Sekundärstruktur wurde die Probe über den temperierbaren Probenhalter des Spektropolarimeters mit 1 °C/min aufgeheizt bzw. abgekühlt und die Elliptizität  $\theta$  bei 218 nm in

1 °C-Schritten aufgezeichnet. Die gemessene Elliptizität wurde anschließend mit Formel 2-2 in die Molare Elliptizität pro Aminosäurerest,  $[\theta]_{MRW}$ , umgerechnet.

**Formel 2-2**

$$[\theta]_{MRW} = \frac{\theta \times 0,1 \times M}{n \times c \times d}$$

$[\theta]_{MRW}$ : mittlere molare Elliptizität pro Aminosäurerest  
 $\theta$ : gemessene Elliptizität [mdeg]  
 $M$ : Molmasse des Proteins [g/mol]  
 $n$ : Anzahl der Peptidbindungen  
 $c$ : Proteinkonzentration [g/l]  
 $d$ : Schichtdicke [cm] ( $\triangleq$  Weglänge des Lichtes durch die Probe)

Durch die Umrechnung in  $[\theta]_{MRW}$  werden CD-Spektren verschiedener Proteine bzw. des gleichen Proteins bei verschiedenen Konzentrationen und Lösungsmittelbedingungen untereinander vergleichbar.

Zur Analyse eines möglichen Einflusses von SDS-Mizellen, SUVs (vgl. Kapitel 2.10) und hydrophober Polystyrolkugeln (PSK, vgl. Kapitel 2.11) auf die Sekundärstruktur von cwCT(F34W) wurden Fern-UV-CD-Spektren von frisch in 5 mM MOPS, pH 7,5, gelöstem Protein (Konzentration  $\sim 15 \mu\text{M}$ ) aufgenommen. Anschließend wurden Fern-UV-CD-Spektren von Ansätzen mit gleicher cwCT(F34W)-Konzentration und entweder  $1,12 \mu\text{M}$  PSK (Konzentrationsberechnung vgl. Kapitel 2.11), 1–50 mM SDS oder 1,45–120,48 nM SUVs ( $\triangleq 0,06$ –5 mM PL-Monomer, Konzentrationsberechnung vgl. Kapitel 2.10) bei verschiedenen Natriumchloridkonzentrationen aufgenommen. Als Kontrollproteine wurden die beiden artifiziellen Spinnenseidenproteine eADF4(C16) (abgeleitet vom *Araneus diadematus* Fibroin 4)<sup>[148]</sup> und Sfl (abgeleitet vom *Nephila clavipes* Flagelliform-Seidenprotein)<sup>[149]</sup> sowie BSA<sup>[186]</sup> (Fraktion V, Sigma-Aldrich) verwendet. Bei den beiden artifiziellen Spinnenseidenproteinen handelte es sich wie beim *flank*-Protein cwCT(F34W) um repetitive, glycinreiche und intrinsisch entfaltete Proteine, deren natürliche Pendants ebenfalls Teil extrakorporaler Proteinfasern sind. BSA wurde als nativ gefaltetes, sehr gut charakterisiertes, extrazelluläres Protein verwendet. Während BSA vor Verwendung direkt in 10 mM Tris/HCl, pH 7,5, bei RT gelöst wurde, wurden eADF4(C16) und Sfl zunächst in 6 M GuaSCN denaturiert und anschließend 2 x 2–3 h bei RT gegen 10 mM Tris/HCl, pH 7,5, dialysiert. Die Konzentration der Kontrollproteine

betrug 2,1  $\mu\text{M}$  für eADF4(C16), 20,9  $\mu\text{M}$  für Sfl und 3,6  $\mu\text{M}$  für BSA, um zu gewährleisten, dass die „Konzentration“ der Peptidbindung ähnlich der bei Aufnahme der Fern-UV CD-Spektren von cwCT(F34W) in gepufferter Lösung vorliegenden war.

Die Zugabe der gelösten Proteine erfolgte immer unmittelbar vor der jeweiligen Messung. Die Fern-UV-CD-Spektren wurden nach kurzem Durchmischen (durch Pipettieren) der Ansätze mit einer Schrittweite von 0,2 nm bei einer Scangeschwindigkeit von 20 nm/min und einer Integrationszeit von 4 s aufgezeichnet und die Werte abschließend ebenfalls in  $[\theta]_{\text{MRW}}$  umgerechnet.

### 2.8.4 Fluoreszenzspektroskopie

Fluoreszenzspektroskopische Analysen wurden in dieser Arbeit dazu verwendet, um Aussagen über die physikalisch-chemische Umgebung des in cwCT(F34W) enthaltenen Tryptophanrests, sowohl in wässrigen Lösungen als auch nach Interaktion des Proteins mit SUV, treffen zu können.

Elektronen eines Moleküls können durch Absorption elektromagnetischer Strahlung aus ihrem elektronischen Grundzustand in einen angeregten Zustand angehoben werden. Bei der Rückkehr in den Grundzustand wird die absorbierte Energie entweder in Wärme umgewandelt oder strahlungsfrei bzw. in Form elektromagnetischer Strahlung abgegeben. Letztgenannter Prozess wird auch als Fluoreszenz und die dafür verantwortlichen chemischen Gruppen als Fluorophore bezeichnet<sup>[187]</sup>. Die durch Fluoreszenz abgegebene elektromagnetische Strahlung hängt dabei von der Schwingungsunterstruktur des elektronischen Grundzustands der Fluorophore ab und erzeugt ein für das jeweilige Molekül charakteristisches Fluoreszenzemissionsspektrum. Änderungen in der direkten Umgebung der Fluorophore, die innerhalb der Lebensdauer des angeregten Zustands (1–100 ns) auftreten und die elektronische Verteilung beeinflussen<sup>[187]</sup>, drücken sich deshalb in Änderungen des Fluoreszenzemissionsspektrums aus.

Intrinsische Fluorophore von Proteinen sind die Seitenketten der aromatischen Aminosäuren Tryptophan, Tyrosin und Phenylalanin. Diese zeigen jedoch starke Unterschiede hinsichtlich ihrer Fluoreszenzintensitäten<sup>[188]</sup> und fluoreszenzspektroskopische Analysen von Proteinen nutzen in erster Linie die Fluoreszenzeigenschaften des Tryptophans. Dessen Seitenkette absorbiert elektromagnetische Strahlung der Wellenlänge 260–290 nm

(mit einem Absorptionsmaximum bei 280 nm) und zeigt in wässriger Lösung eine  $\sim 60$  nm breite Fluoreszenzemissionsbande, deren Maximum bei  $\sim 350$  nm liegt<sup>[189]</sup>. Zur Aufnahme der Tryptophan-Fluoreszenzemissionsspektren von cwCT(F34W) wurden 8,4  $\mu$ M Lösungen des in 5 mM MOPS, pH 7,5, gelösten Proteins verwendet. Zusätzlich wurden Fluoreszenzemissionsspektren in Gegenwart von  $\sim 0,125$   $\mu$ M DSPG-SUVs ( $\cong 5$  mM DSPG-Monomer, vgl. Kapitel 2.10) aufgezeichnet, um Auswirkungen der cwCT(F34W)/DSPG-SUV-Interaktion auf die chemische Umgebung des Tryptophans zu analysieren. Zur Aufzeichnung der jeweiligen Spektren bei RT wurde die Tryptophanfluoreszenz in einem Jasco FP-6500 Spektrofluorimeter (Jasco) bei einer Wellenlänge von 280 nm (Spaltbreite 10 nm) angeregt und die Emission von 300–450 nm (Spaltbreite 5 nm) mit 0,2 nm Schrittweite, einer Scan-Geschwindigkeit von 50 nm/min und einer Integrationszeit von 1 s aufgezeichnet.

## 2.9 MALDI-TOF Massenspektrometrie

Die matrixunterstützte Laser-Desorption/Ionisation (MALDI, von engl. *matrix-assisted laser desorption/ionization*) mit Kopplung an einen Flugzeit-Massendetektor (TOF-Detektor, von engl. *time of flight*) eignet sich vor allem zur exakten Bestimmung der Masse von Peptiden und Proteinen (sowie anderer großer Biomoleküle und Molekülkomplexe), ohne diese zuvor fragmentieren zu müssen. Das zu analysierende Molekül - der Analyt - wird zusammen mit den Molekülen der verwendeten Matrix kristallisiert. Die Moleküle der Matrix besitzen die Eigenschaft, elektromagnetische Strahlung im Wellenlängenbereich des zur Ionisation verwendeten Lasers (in der Regel ein Stickstofflaser mit einer Wellenlänge von 337 nm) zu absorbieren<sup>[190]</sup>. Dies führt zu einer schlagartigen Sublimation der Matrixmoleküle, wodurch auch Analytmoleküle mit in die Gasphase gelangen. Dort führen v. a. Protonenübertragungen zu einer Ionisation des Analyten. Die gebildeten Ionen werden in einem elektrischen Feld beschleunigt und deren Flugzeit durch den TOF-Detektor ermittelt. Da die Flugzeit proportional dem Quadrat des Quotienten aus Masse und Ladung eines ionisierten Moleküls ist, kann über deren Messung die Masse des Analyten exakt bestimmt werden.

Die Vorbereitung der Proben für MALDI-TOF Analysen wurde von Dr. Martin Humenik durchgeführt. Die Probenvorbereitung umfasste die Entsalzung der Proteinproben mit

Hilfe kommerziell erhältlicher C4 und C18 ZipTip-Pipettenspitzen (Merck Millipore) sowie das Mischen der entsalzten Proben 1:1 mit Matrixlösung (20 mg/ml Sinapinsäure in 90 % (v/v) Acetonitril). Die MALDI-TOF Massenspektren wurden von Christopher Synatschke am Lehrstuhl für Makromolekulare Chemie II der Universität Bayreuth an einem Reflex III-Massenspektrometer (Bruker) detektiert. Die Analyse und Auswertung der Massenspektren erfolgte mit Hilfe der Software mMass<sup>[150]</sup>.

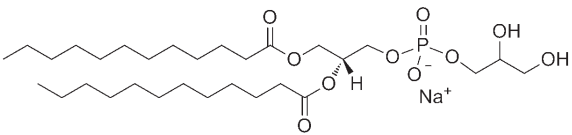
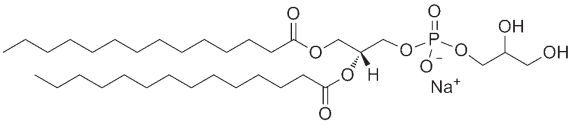
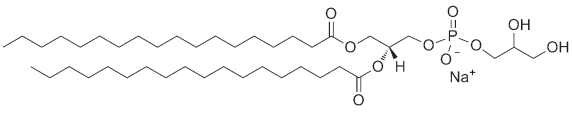
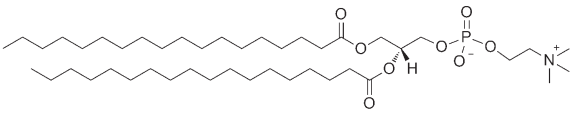
### 2.10 Herstellung von SDS-Mizellen und Emulsionen kleiner unilamellarer Lipid-vesikel (SUV)

SDS wird in der Biochemie neben seiner klassischen Rolle als Proteindenaturierungsmittel auch als stark vereinfachtes Modellsystem zur Simulation von Lipidvesikeln eingesetzt<sup>[191]</sup>, um Aufschlüsse über mögliche Interaktionen von Proteinen mit biologischen Lipidmembranen zu gewinnen. SDS selbst bildet oberhalb seiner kritischen Mizellbildungskonzentration (CMC, von engl. *critical micelle concentration*) spontan Mizellen aus. Die CMC für SDS ist nicht konstant, sondern liegt je nach Temperatur, pH-Wert, Ionenstärke der Lösung oder auch dem Vorhandensein weiterer Moleküle zwischen 0,14–8 mM<sup>[191,192]</sup>. Zur Untersuchung eines möglichen Einflusses von SDS-Mizellen auf die Sekundärstruktur des rekombinanten *flank*-Proteins cwCT(F34W) wurde eine SDS-Mizellen-Stammlösung aus 100 mM SDS in MQ-Wasser hergestellt und in Konzentrationen von 1–50 mM in den entsprechenden CD-spektroskopischen Analysen (vgl. Kapitel 2.8.3) verwendet.

Darüber hinaus wurden als komplexeres Modellsystem für biologische Lipiddoppelmembranen SUV hergestellt. Als Ausgangssubstanzen wurden die Phospholipide (PL) DLPG, DMPG, DSPG und DSPC (alle von Avanti<sup>®</sup> Polar Lipids) verwendet (Tab. 2-5). Zur Herstellung von SUV-Emulsionen wurden ~ 10 mmol des PL eingewogen, 1,5 ml Trichlormethan zugegeben und das PL bei Raumtemperatur durch regelmäßig leichtes Schwenken gelöst. Danach wurde das Chloroform zunächst 1 h unter Vakuum in einem 20 °C warmen Wasserbad und anschließend für weitere 3 h im 40 °C warmen Wasserbad abgezogen. Zum getrockneten Lipidkuchen wurde das Volumen MQ-Wasser gegeben, das zum Einstellen einer PL-Endkonzentration von 10 mM benötigt wurde, und, um ein Quellen des PL-Kuchens zu ermöglichen, für 30–60 min (ohne Vakuum) bei 70 °C im

Wasserbad weiter rotiert. Durch abschließende, zehninütige Ultraschallbehandlung des gequollenen PL im 70 °C warmen Ultraschallbad wurde die für weitere Experimente genutzte SUV-Emulsion erzeugt.

**Tab. 2-5. Übersicht der verwendeten artifiziellen Phospholipide.** Die Tabelle zeigt für jedes der verwendeten artifiziellen Phospholipide DLPG, DMPG, DSPG und DSPC deren Strukturformel, das Molekulargewicht ( $M_w$ ), die kritische Mizellbildungskonzentration (CMC) und die Ladung der polaren Kopfgruppe. n.b.: nicht bekannt.

Phospholipid:	Strukturformel:	$M_w$	CMC [mmol/l]	Ladung der Kopfgruppe:
DLPG		632,729	0,130	negativ
DMPG		688,845	0,011	negativ
DSPG		801,058	n.b.	negativ
DSPC		790,145	n.b.	zwitter-ionisch

Hergestellte SUV-Emulsionen wurden bei 4 °C nicht länger als 7 Tage gelagert, um oxidative Abbauprozesse der Phospholipide zu vermeiden. Zudem wurden gelagerte SUV-Emulsionen direkt vor Verwendung erneut für 10 min im 70 °C warmen Wasserbad ultraschallbehandelt, um eine möglichst hohe Homogenität der SUV-Emulsionen zu gewährleisten.

Da die Herstellerfirma Avanti Polar Lipids weder Informationen bezüglich der CMC noch der Anzahl der PL-Moleküle pro SUV (Aggregationsnummer) zur Verfügung stellte, wurde letztere näherungsweise mit vereinfachten Annahmen berechnet. Ausgehend von einem durch Transmissionselektronenmikroskopie (TEM) bestimmten durchschnittlichen Vesikeldurchmesser von 75 nm, der Annahme von kugelförmigen Vesikel sowie einer durchschnittlichen Lipiddoppelmembrandicke von 6 nm<sup>[193]</sup> wurde die Oberfläche einer

Kugel mit 75 nm und einer Kugel mit 63 nm Durchmesser (die Differenz der Durchmesser entspricht der zweifachen Doppelmembrandicke von 6 nm) berechnet.

**Formel 2-3**  $A_g = 4 \times \pi \times (r_1^2 + r_2^2)$

$A_g$ : Gesamtfläche  
 $r_1$ : Radius Kugel 1  
 $r_2$ : Radius Kugel 2

Die durch Formel 2-3 ermittelte Gesamtfläche von  $\sim 30140 \text{ nm}^2$  wurde anschließend durch einen näherungsweise angenommenen PL-Monomerflächenbedarf von  $72 \text{ Å}^2$  geteilt<sup>[194]</sup>. Auf diesem Weg wurde als Aggregationsnummer der DSPG-SUV 41861 berechnet. Somit wurde für alle Versuche, bei denen DSPG-SUV beteiligt waren, angenommen, dass ein durchschnittliches DSPG-SUV  $\sim 42000$  Moleküle umfasste und 10 mM DSPG somit einer „Vesikelkonzentration“ von  $\sim 0,240 \text{ μM}$  entsprach.

### 2.11 Präparation von Polystyrol-Nanokügelchen

Im Rahmen dieser Arbeit wurden kommerziell erhältliche Polystyrolkügelchen (PSK, Handelsname: *3000 Series Nanosphere™ Size Standards*, Thermo Fisher Scientific) mit Durchmessern von  $30 \pm 1 \text{ nm}$  zur Simulation gekrümmter hydrophober Oberflächen als Negativkontrolle bei der Analyse der Interaktionen des rekombinanten *flank*-Proteins cwCT(F34W) mit SUV verwendet. Ausgehend von der vom Hersteller angegebenen Dichte, dem durchschnittlichen Volumen eines PSK sowie der Massenkonzentration der PSK-Emulsion konnte nach Formel 2-4 die Stoffmengenkonzentration (in mol/l) berechnet werden.

**Formel 2-4** 
$$n = \frac{10g}{\rho \times V \times N_A}$$

$n$ : Stoffmenge [mol] (in 1 l PSK-Emulsion)  
 $\rho$ : Dichte eines PSK  
 $V$ : Volumen eines PSK  
 $N_A$ : Avogadrokonstante

Unter Berücksichtigung der Herstellerangaben betrug die Konzentration der PSK somit  $\sim 1,12 \text{ μM}$ .



Die PSK-Lösung wurde zunächst 45 min bei 186 000 x g zentrifugiert, um die PSK zu sedimentieren. Anschließend wurden die PSK durch Ultraschallbehandlung mit dem Bandelin SONOPULS GM 3200 Ultraschallgenerator (Bandelin) und der Ultraschallsonde MS73 für 1 min bei 20 % der Maximalamplitude in 1 ml 10 mM MOPS, pH 7,5, 150 mM Natriumchlorid, resuspendiert. Nach Resuspension wurden die PSK auf die o. g. Weise noch zwei weitere Mal gewaschen. Abschließend wurde die UV-Absorption im Wellenlängenbereich < 260 nm gemessen, um auszuschließen, dass keine die späteren Fern-UV-CD-Messungen (vgl. Kapitel 2.8.3) störenden Verunreinigungen mehr vorhanden waren.



## 3 Ergebnisse

### 3.1 Bioinformatische Charakterisierung der flankierenden Regionen der preCols

#### 3.1.1 Vergleich der cDNA-Bank preCol-Sequenzen mit bekannten Sequenzen

Ausgangspunkt für die in dieser Arbeit verwendeten Proteine, die den Sequenzen der flankierenden Regionen von *M. galloprovincialis* preCol-D, preCol-P und preCol-NG nachempfunden sind, war die am Lehrstuhl Biomaterialien verwendete cDNA-Datenbank (Mfl-1, von engl. *mus*sel *foot* *library*) des Muschelfußes von *M. galloprovincialis*<sup>[125]</sup>. Zunächst wurden die in der cDNA-Datenbank enthaltenen Nukleotidsequenzen der drei preCols translatiert und mittels LALIGN<sup>[155]</sup> ([http://embnet.vital-it.ch/software/LALIGN\\_form.html](http://embnet.vital-it.ch/software/LALIGN_form.html)) mit bereits publizierten *M. galloprovincialis* Aminosäuresequenzen verglichen<sup>[109]</sup>.

Die lokalen, paarweisen Aminosäurealignments der flankierenden Regionen belegten, dass es sich bei den aus der Mfl-1 stammenden preCol-Sequenzen in Bezug auf die bereits veröffentlichten Sequenzen um Proteinisoformen handelte (Abb. 3-1). Im Bereich der flankierenden Regionen ließen sich in allen Fällen kleinere Unterschiede in den Aminosäuresequenzen feststellen (Abb. 3-1 A–C, gelb markierte Aminosäuren). Dabei zeigten die carboxyterminal flankierenden Regionen der preCols P und D eine VGGG bzw. eine GGA-Insertion. Die im Fall von preCol-P vorliegende VGGG-Insertion stellte eine Erweiterung des durch elastinähnliche  $X^1X^2X^3$ PG-Motive gekennzeichneten Blocks dar (Abb. 3-1 A). Gleiches galt für die carboxyterminal flankierende Region von preCol-D, bei der die Insertion der GGA-Sequenz zur Erweiterung eines seidenproteintypischen GGX-Blocks führte (Abb. 3-1 B). In beiden Fällen blieben sowohl der repetitive Charakter als auch die charakterisierenden Motive erhalten. Die aminoterminalen flankierenden Regionen der preCols P und D zeigten nur vereinzelte Aminosäuresubstitutionen, die den jeweiligen Sequenzcharakter der Konsensusmotive ebenfalls nicht beeinflussten (Abb. 3-1 A & B).

### 3 Ergebnisse

#### preCol-P

		Aminotermminus	
Mf1-1	1	GGFGGGIGGIGPGGSVGGGIGGIGGGVGGVGGIGGGVGGPGG	IGGIGIGPGFGGGFGPGSSAS
AF448525	129	.....	N.....
Mf1-1	61	GS	SGSAFGGPGGSSASANAARANANGGGGFG
AF448525	189	F.....	.....
		Carboxyterminus	
Mf1-1	1	FPGFGGFGGFGGASANAGSSANANAGPGGSAGAGSSSGANANAGGFPFGGPFGGVGGGP	
AF448525	674	.....	.....
Mf1-1	61	GGAGGPGGVGGGPGGVGGGVGGGPGGVGGG	VGGGPGGAGGLGGFGGFGGGSSAGASSSSGS
AF448525	734	.....	-----
Mf1-1	121	ASASSGGPFVGLNVGPGGGIGGGSASASAASRA	
AF448525	790	.....	.....

#### preCol-D

		Aminotermminus	
Mf1-1	1	GGASAAAAAARASAGGLGGLGGF	GSAAANAAAAANAGAGFGGFGGFGGF
AF448526	89	.....	V.....L.....
Mf1-1	61	GPGGPGGPGGPGGPGMPGGPGGPGSGPTGGP	
AF448526	149	.....	.....
		Carboxyterminus	
Mf1-1	1	GGLGGAGPGGLGGLGGGLGGLGGGLGGLGGAGGLGGGLGGLGGAG	GG
AF448526	732	.....	--
Mf1-1	61	AGGAGGAGGAGATAAAAAQAAAAAGAGNGGAAQAAAQAAASAAANSGLGAGAARAAA	SA
AF448526	790	.....	.....
Mf1-1	121	AARATVAGHGSGTAAAAANAAAQ	
AF448526	849	.....	.....

#### preCol-NG

		Aminotermminus	
Mf1-1	1	AAAAHAAVANALGGGGGGSASAAAAAARALGGGGGSASAAAAAARALGGGGGAS	
AF448524	58	.....	.....
Mf1-1	61	AAARAAAAASALGGGGG	-FGGLGG---LGGGAGGLGGLGGGLGGLGGGGVGGGGVGGGG
AF448524	176	.....G.....GLGG.....F.....P.....LGG--	
Mf1-1	116	VGPGLLGGGGGPGSAGAFGNGNAAAGPGGAGASASAGAFATGGGGFPLPG	
AF448524	119	.....	.....
		Carboxyterminus	
Mf1-1	1	GAGE	GGAGGAGGAGGAGGLGGGAGGAGGLGGGFGLGGGAGGLGGLGGGAGGAGGVGGL
AF448524	715	....A.T.....P.....EPR-	.....
Mf1-1	61	GGGVGGLGGVGLGFGGAGASAGAGANAGAGGAGGSASA	
AF448524	774	.....	.....

**Abb. 3-1. Alignment der Aminosäuresequenzen der flankierenden Regionen der *M. galloprovincialis* preCols.** Die aus der Mf1-1 cDNA-Bank des Lehrstuhls Biomaterialien stammenden Aminosäuresequenzen für die flankierenden Regionen der *M. galloprovincialis* preCols im Vergleich zu den von Lucas *et al.* publizierten<sup>[109]</sup> Sequenzen. Für die bereits publizierten Sequenzen ist die GenBank Zugangsnummer angegeben. Sequenzunterschiede sind gelb markiert.

Die größten Sequenzunterschiede zwischen den von Lucas *et al.* publizierten<sup>[109]</sup> und aus der Mf1-1 erhaltenen Sequenzen waren bei preCol-NG erkennbar. Durch Insertionen sowie Deletionen unterschieden sich die aminoterminal flankierenden Regionen in Anzahl und Länge der charakterisierenden (XG<sub>m</sub>)<sub>n</sub>-Bereiche und zeigten v. a. auf der

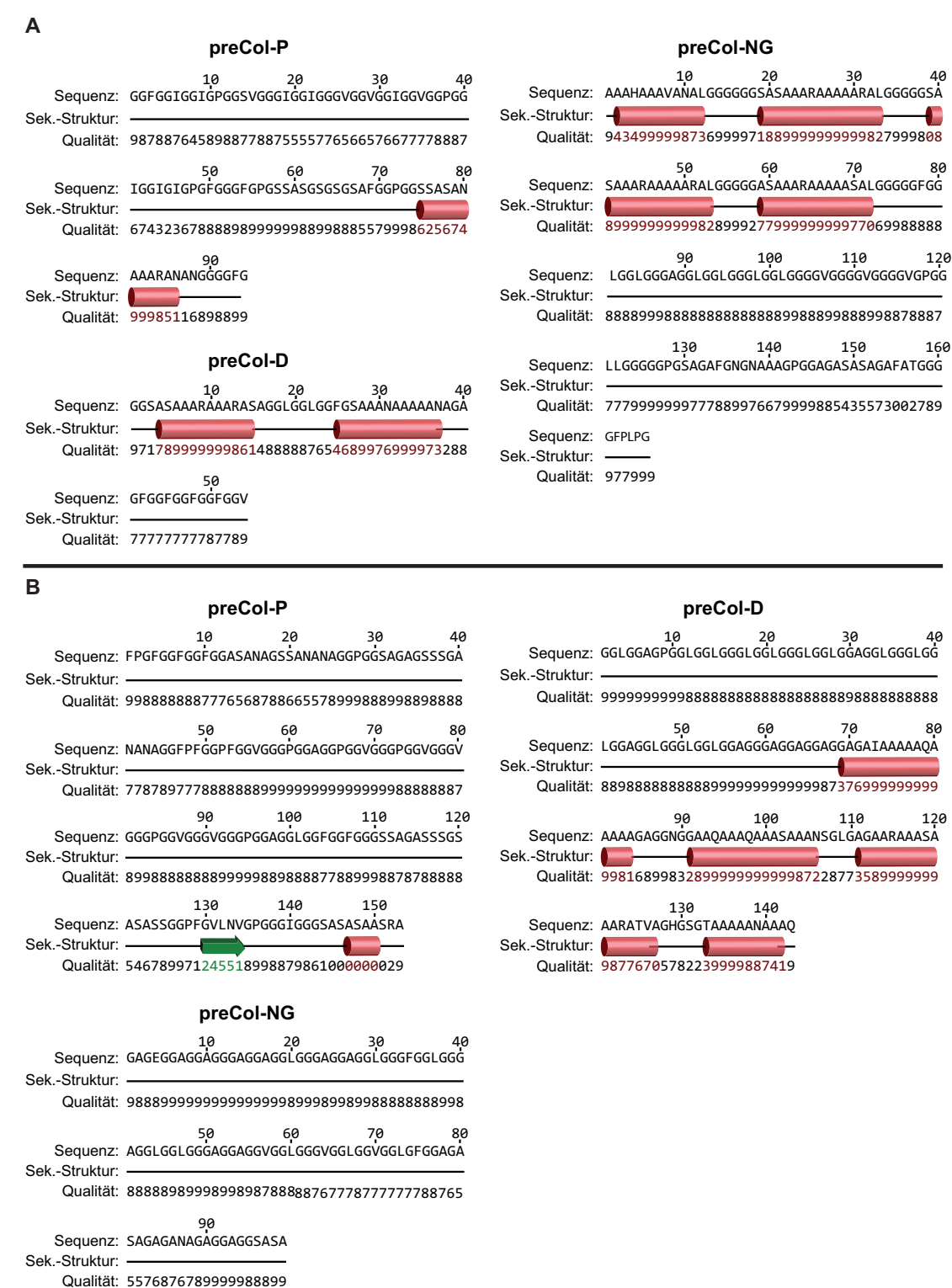
X-Position der Konsensusmotive Aminosäuresubstitutionen (Abb. 3-1 C). Bei den Substitutionen handelte es sich jedoch ausschließlich um Austausche von häufig vorkommenden durch andere innerhalb der Konsensusmotive ebenfalls häufig vorkommenden Aminosäuren. Die carboxyterminal flankierende Region von preCol-NG war gegenüber der veröffentlichten Sequenz ausschließlich durch Aminosäuresubstitutionen gekennzeichnet. Dass die Mfl-1 basierte Sequenz – im Gegensatz zu der von Lucas *et al.*<sup>[109]</sup> publizierten – eine strikte Einhaltung der Konsensusmotive aufweist, bestätigt die schon bekannte Tatsache<sup>[125]</sup>, dass das bei letzterer identifizierte EPR-Aminosäuremotiv wohl auf eine Leserasterverschiebung während der Sequenzierung zurückzuführen ist.

Die Alignments zeigten, dass eine rekombinante Herstellung von Proteinen, die den flankierenden Regionen der drei preCols entsprechen, auf Grundlage der entsprechenden Mfl-1 Sequenzen durchgeführt werden kann.

### 3.1.2 Sekundärstrukturvorhersage der flankierenden Regionen der preCols

In Vorbereitung der rekombinanten Herstellung der flankierenden Regionen der preCols als eigenständige Proteine wurde zunächst eine bioinformatische Vorhersage der Sekundärstruktur mittels des auf einem neuronalen Netzwerk basierenden Programms PSIPRED<sup>[195]</sup> durchgeführt.

Die unabhängig von der Volllängensequenz der jeweiligen preCols durchgeführte PSIPRED-Analyse der Aminosäuresequenzen der flankierenden Regionen von preCol-D, -P und -NG deutete darauf hin, dass, außer in der carboxyterminal flankierenden Region von preCol-NG, in allen flankierenden Regionen die Ausbildung von Sekundärstrukturelementen wahrscheinlich ist (Abb. 3-2). Für die aminoterminalen Regionen aller drei untersuchten preCols wurde das Auftreten  $\alpha$ -helikaler Sekundärstruktur vorhergesagt (Abb. 3-2 A). Dabei viel auf, dass es sich bei den Aminosäuresequenzen, für die  $\alpha$ -Helices vorhergesagt wurde, um vereinzelt durch Aminosäuren mit großen Seitenketten unterbrochene Polyalaninbereiche handelte.



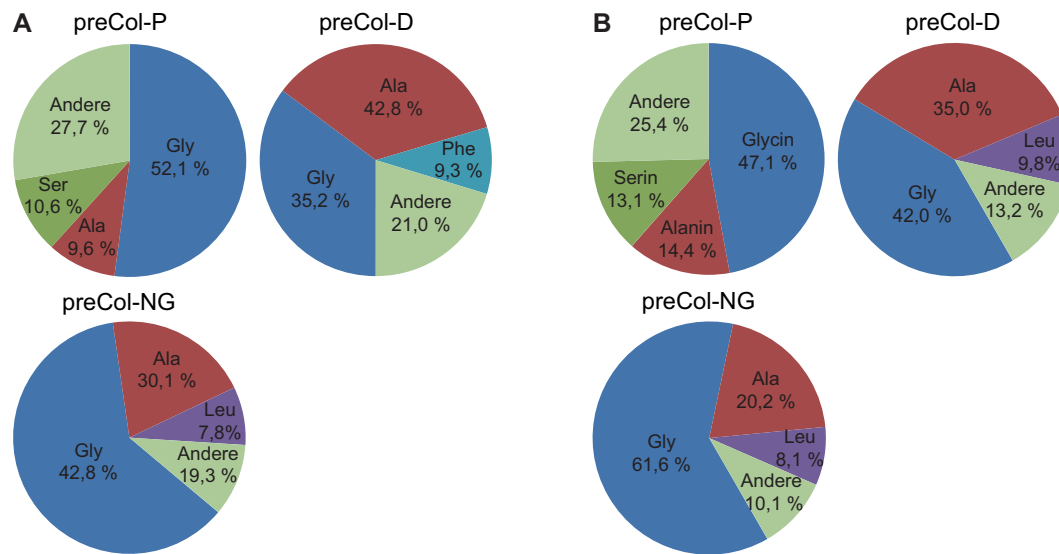
**Abb. 3-2. PSIPRED Sekundärstrukturvorhersage der flankierenden Regionen von *M. galloprovincialis* preCol-P, preCol-D und preCol-NG. (A)** Sekundärstrukturvorhersage für die aminoterminal flankierenden Bereiche von preCol-P, -D und -NG. **(B)** Sekundärstrukturvorhersage für die carboxyterminal flankierenden Bereiche von preCol-P, -D und -NG. Die Sekundärstrukturvorhersagen unterscheiden unstrukturierte Bereiche (schwarze Linie) von  $\alpha$ -Helices (rote Zylinder) und  $\beta$ -Faltblattstrukturen (grüne Pfeile). „Qualität“ beschreibt die Zuverlässigkeit der Sekundärstrukturvorhersage und kann Werte zwischen 0 (höchst unwahrscheinlich) und 9 (höchst wahrscheinlich) annehmen.

Im Fall der carboxyterminal flankierenden Regionen wurden nur für die des preCol-D nennenswerte Sekundärstrukturanteile vorhergesagt (Abb. 3-2 B). Dabei handelte es sich um vier Bereiche, die ebenfalls mit großen Aminosäuren substituierte Polyalaninsequenzen umfassen und zur Ausbildung einer  $\alpha$ -helikalen Konformation neigen. Für die carboxyterminal flankierende Region von preCol-NG wurde keinerlei, sowie für die flankierende Region von preCol-P lediglich ein kurzer  $\alpha$ -helikaler Abschnitt und ein kurzer  $\beta$ -Faltblattbereich vorhergesagt. Wie die Werte der Vorhersagequalität belegten, ist die Existenz des  $\alpha$ -helikalen Bereiches jedoch höchst unwahrscheinlich. Für sämtliche andere Bereiche mit ungeordneter sowie  $\alpha$ -helikaler Konformation deuteten die Qualitätswerte auf eine hohe Zuverlässigkeit der Sekundärstrukturvorhersage hin.

### **3.1.3 Sequenzanalyse der flankierenden Regionen bezüglich Aminosäurezusammensetzung und Codonverwendung**

#### *3.1.3.1 Aminosäurezusammensetzung*

Die stark repetitiven Aminosäuresequenzen der flankierenden Regionen der preCols sind dadurch charakterisiert, dass sie nur wenige verschiedene Aminosäuren umfassen, was jedoch zu einem überproportional hohen Anteil dieser Aminosäuren beiträgt. Sowohl in den amino- (Abb. 3-3 A) wie carboxyterminal (Abb. 3-3 B) flankierenden Regionen der drei *M. galloprovincialis* preCols macht die Aminosäure Glycin zwischen ein und zwei Drittel aller Aminosäuren aus. Darüber hinaus ist, von der aminoterminal flankierenden Region von preCol-P abgesehen (Abb. 3-3 A), Alanin mit einem Anteil, der zwischen 14,4 % und 42,8 % ausmacht, immer die am zweithäufigsten vorkommende Aminosäure. Dies führt dazu, dass alleine die beiden genannten Aminosäuren in allen Fällen mindestens 52 % aller in den flankierenden Regionen vorkommender Aminosäuren stellen. Darüber hinaus ist in beiden flankierenden Regionen von preCol-P nur noch Serin mit einem Anteil von über 10 % vertreten.



**Abb. 3-3. Aminosäureanteile der flankierenden Regionen der *M. galloprovincialis* preCols.** Dargestellt sind die Anteile der drei jeweils am häufigsten in den amino- (A) und carboxyterminal (B) flankierenden Sequenzen der preCols vorkommenden Aminosäuren.

#### 3.1.3.2 Vergleich der Codonverwendung von *M. galloprovincialis* und *E. coli*

Aufgrund der Degeneration des genetischen Codes existieren für die genetische Codierung jeder der kanonischen Aminosäuren jeweils zwischen ein und sechs verschiedene Basentriplets (sog. synonyme Codons). Dies bedingt, dass verschiedene Lebewesen im Lauf der Evolution auch Unterschiede bei der Verwendungshäufigkeit synonymen Codons entwickelten, was Einfluss auf die heterologe Produktion transgener Proteine nehmen kann. Für die rekombinante Herstellung der flankierenden Regionen der *M. galloprovincialis* preCols in *E. coli* wurden die entsprechenden Mfl-1-Nukleotidsequenzen mit Hilfe des *Graphical Codon Usage Analyzer 2.0*<sup>[158]</sup> hinsichtlich der *E. coli*-Codonverwendung analysiert. Tabelle 3-1 zeigt das Ergebnis für Codons der Aminosäuren, die in den Sequenzen der flankierenden Regionen der preCols am häufigsten vertreten sind, da ausgeprägte Unterschiede in der Codonverwendung in diesen Fällen zu Problemen bei der heterologen Produktion der entsprechenden Proteine in *E. coli* führen können.

Ein Vergleich der Codonverwendung von *M. galloprovincialis* für die Aminosäure Glycin, welche mit Anteilen von 35,2–61,1 % die am häufigsten in den flankierenden Regionen der preCols vorkommende Aminosäure darstellt (Abb. 3-3 A & B), mit der



Codonverwendung von *E. coli* zeigte, dass das in den preCol-Sequenzen am häufigsten verwendete Codon GGA unter den vier für Glycin existierenden synonymen Codons in *E. coli* Gensequenzen nur das dritthäufigste ist. Hingegen wird das in den Sequenzen der flankierenden Regionen aller preCols am zweithäufigsten genutzte Glycincodon GGT von *E. coli* am häufigsten zur Glycincodierung verwendet (Tab. 3-1).

**Tab. 3-1. Vergleich der Codonverwendung von *M. galloprovincialis* und *E. coli*.** Die Codons der drei in der jeweiligen flankierenden Region der *M. galloprovincialis* preCols am häufigsten vorkommenden Aminosäuren sind der Codonverwendung von *E. coli* gegenübergestellt. X|Y bedeutet, dass das entsprechende Codon am X-häufigsten in der entsprechenden *M. galloprovincialis*-Sequenz vorkommt und von *E. coli* am Y-häufigsten verwendet wird. Rot: in der entsprechenden *M. galloprovincialis*-Sequenz am häufigsten für die jeweilige Aminosäure vorkommendes Codon; blau: von *E. coli* am häufigsten für die jeweilige Aminosäure verwendetes Codon.

preCol		aminoterminal			carboxyterminal		
		P	D	NG	P	D	NG
Glycin:	GGA	1 3	1 3	1 3	1 3	1 3	1 3
	GGC	2 2	2 2	3 2	3 2	3 2	3 2
	GGG				4 4		
	GGT	2 1	2 1	2 1	2 1	2 1	2 1
Alanin:	GCA	1 1	1 1	1 1	1 1	1 1	1 1
	GCC		2 2	3 2	3 2	2 2	2 2
	GCT	2 4		2 4	2 4	2 4	3 4
Serin:	AGC				2 1	1 1	
	AGT				4 2	3 2	
	TCA				1 2	2 2	
	TCC				5 5		
	TCT				3 2		
Phenylalanin:	TTC		1 2				
	TTT		2 1				
Leucin:	CTA			2 6			
	CTC			2 5		4 6	1 5
	CTG			3 1		1 5	2 1
	CTT			1 3		4 1	2 3
	TTA			3 2		3 3	2 2
	TTG			3 4		2 2	
Isoleucin:	ATC	2 2					
	ATT	1 1					

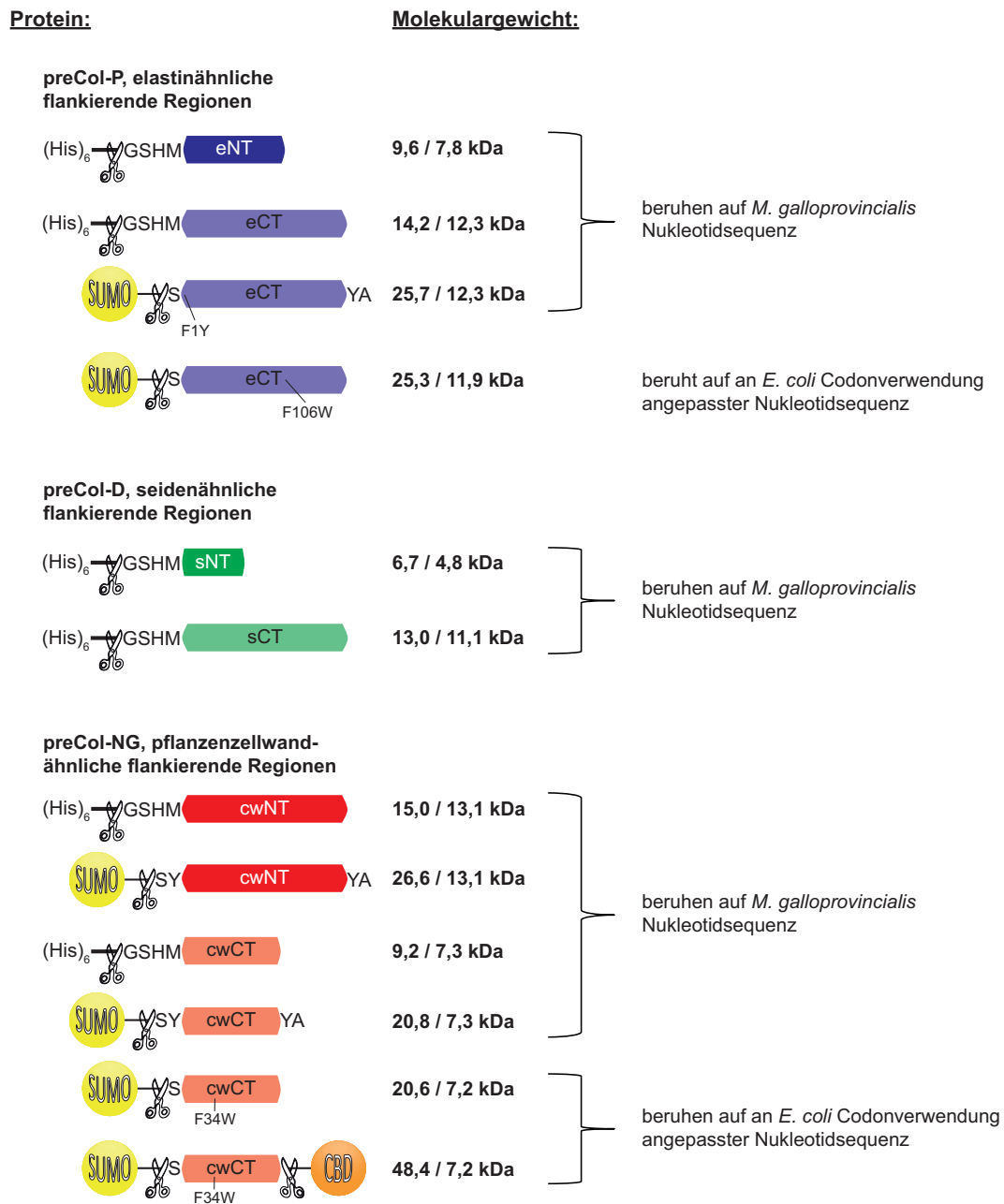
Im Gegensatz zu den für Glycin verwendeten Codons zeigte ein Vergleich der Codons, die für die mit einem Anteil von 14,4–42,8 % (mit Ausnahme der aminoterminal flankierenden Region von preCol-P) am zweithäufigsten in den flankierenden Regionen vorkommende Aminosäure Alanin codieren (Abb. 3-3 A & B), deutlich geringere

Unterschiede hinsichtlich der Codonverwendung in *M. galloprovincialis* und *E. coli*. In beiden Fällen war GCA das am häufigsten zur Codierung von Glycin verwendete Codon. Dem gegenüber ist das Codon GCT, welches mit Ausnahme der aminoterminal flankierenden Region von preCol-D sowie der carboxyterminal flankierenden Region von preCol-NG das nach GCA am zweithäufigsten vorkommende Alanincodon in den entsprechenden preCol-Nukleotidsequenzen ist, in *E. coli*-Nukleotidsequenzen das generell am seltensten (Tab. 3-1) verwendete. Während Glycin und Alanin, mit Ausnahme der aminoterminal flankierenden Region von preCol-P, generell die beiden am häufigsten auftretenden Aminosäuren sind, unterscheidet sich die am dritthäufigsten vorkommende je nach preCol-Molekül (Abb. 3-3). Vor allem im Fall der Phenylalanin- (aminoterminal flankierende Region von preCol-D) und der Leucincodons (carboxyterminal flankierende Regionen von preCol-D und -NG) konnten, ähnlich wie bei den Glycincodons, erneut Unterschiede der Codonverwendung zwischen *M. galloprovincialis* und *E. coli* festgestellt werden (Abb. 3-3).

### 3.2 Herstellung rekombinanter *M. galloprovincialis* flank-Proteine

Auf Grundlage der durchgeführten bioinformatischen Analysen (vgl. Kapitel 3.1) und der aus der Mfl-1 bekannten Nukleotidsequenzen wurden verschiedene, im Folgenden als *flank*-Proteine bezeichnete, Proteine rekombinant in *E. coli* hergestellt (Abb. 3-4).

Neben den auf den entsprechenden *M. galloprovincialis* Nukleotidsequenzen beruhenden *flank*-Proteinen wurden auch auf hinsichtlich der *E. coli* Codonverwendung optimierten Nukleotidsequenzen beruhende *flank*-Proteine produziert. Zur genaueren Beschreibung einzelner *flank*-Proteine werden diese entsprechend der Trivialbezeichnung der flankierenden Regionen der preCols (vgl. Kapitel 1.3.2) als elastinähnlich (eNT und eCT aus preCol-P), seidенähnlich (sNT und sCT aus preCol-D) oder zellwandähnlich (von engl. „cell wall-like“, cwNT und cwCT aus preCol-NG) bezeichnet. Darüber hinaus lassen die letzten beiden Buchstaben erkennen, ob es sich um die entsprechende aminoterminal (NT) oder carboxyterminal (CT) flankierende Region handelt. Zur strukturellen und funktionellen Charakterisierung wurden zusätzlich noch Tyrosin- sowie Tryptophanmutanten mit verbesserten UV-spektroskopischen Eigenschaften erzeugt (Proteinübersicht vgl. auch Kapitel 11.1).

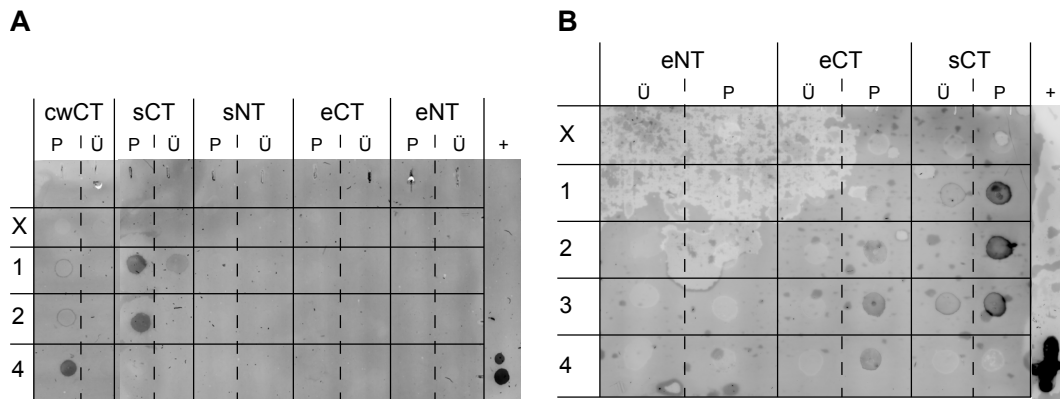


**Abb. 3-4. Übersicht aller rekombinant in *E. coli* hergestellten *flank*-Proteine.** Alle Proteine wurden in Form von Fusionsproteinen mit verschiedenen absplaltbaren *tags* hergestellt. Die zugrunde liegende Nukleotidsequenz wurde entweder direkt aus Mfl-1 amplifiziert oder war zuvor an die Codonverwendung von *E. coli* angepasst worden. Erweiterungen der eigentlichen Aminosäuresequenzen der flankierenden Regionen (blau/violett, grün/hellgrün und rot/rosa hinterlegt) waren entweder durch Klonierung und *tag*-Absplaltungen bedingt oder wurden, im Fall der Tyrosinmutanten, absichtlich erzeugt. Für jedes Protein ist das Molekulargewicht mit und ohne *tags* angegeben. SUMO: SUMO-*tag*; CBD: Chitinbindedomäne (als *tag*); NT: aminoterminal flankierende Region des jeweiligen preCol-Moleküls; CT: carboxyterminal flankierende Region des jeweiligen preCol-Moleküls.

### 3.2.1 Herstellung unveränderter preCol *flank*-Proteine

Ausgehend von den für die flankierenden Regionen der preCols festgelegten Sequenzabschnitten (Abb. 3-1) wurden Primer entworfen, mit denen die entsprechenden Nukleotidsequenzen durch Polymerasekettenreaktion (PCR, von engl.: *polymerase chain reaction*) amplifiziert werden konnten. Die anschließende erfolgreiche Ligation mit dem *E. coli* Expressionsvektor pET28a (Abb. 2-1) sollte die Herstellung der *flank*-Proteine als Fusionsproteine mit spaltbarem, aminoterminalen (His)<sub>6</sub>-tag ermöglichen. Durch Sequenzierungen konnte bestätigt werden, dass die mit dem Expressionsvektor ligierten Nukleotidsequenzen denen der flankierenden Regionen der Mfl-1 codierten preCols entsprachen.

Die Produktion der sechs Fusionsproteine (His)<sub>6</sub>-eNT, -eCT, -sNT, -sCT, -cwNT und -cwCT (Abb. 3-4) wurde in den *E. coli* Expressionsstämmen HMS174(DE3), BL21(DE3), BL21-Gold(DE3) und BL21-CodonPlus(DE3)-RIPL (vgl. Kapitel 2.1.5) bei verschiedenen Temperaturen getestet. Ein Nachweis gebildeter Fusionsproteine in den entsprechenden *E. coli* Zelllysaten gelang weder durch Coomassie- noch durch Silberfärbung nach SDS-PAGE und auch Westernblotanalysen gegen den an jedem *flank*-Protein aminoterminal vorhandenen (His)<sub>6</sub>-tag führten nur zu nicht eindeutigen Ergebnissen, die sich nicht reproduzieren ließen. Um ausschließen zu können, dass ein Nachweis der Proteine nach gelelektrophoretischer Auftrennung der Proben nur aufgrund ihrer geringen Größe nicht möglich war, wurden zudem Dot-Blot Analysen der Zelllysate der Expressionstest durchgeführt. Die besten Resultate hinsichtlich der rekombinanten Produktion konnten mit dem *E. coli*-Stamm BL21-CodonPlus(DE3)-RIPL bei einer Temperatur von 30 °C erzielt werden (Abb. 3-5). Die Dot-Blot Analysen zeigten, dass unter den erwähnten Bedingungen nur die *flank*-Fusionsproteine (His)<sub>6</sub>-eCT, -sCT und -cwCT gebildet werden konnten (Abb. 3-5 A & B). Gebildetes eCT, sCT und cwCT war nach Lyse der Bakterienzellen zudem hauptsächlich in der unlöslichen Fraktion zu finden und nur sCT konnte in Spuren auch in der löslichen Fraktion nachgewiesen werden (Abb. 3-5 B).



**Abb. 3-5: Dot-Blot Analysen der Herstellung von *flank*-Proteinen in *E. coli* BL21-CodonPlus(DE3)-RIPL bei 30 °C. (A)** Immunochemischer Nachweis der aminoterminalen (His)<sub>6</sub>-tags der rekombinant gebildeten *flank*-Fusionsproteine. **(B)** Immunochemischer Nachweis der Proben aus (A) unter Verwendung doppelter Probenmengen. X: Probe vor Induktion; 1–4: Proben 1, 2, 3 und 4 h nach Induktion; Ü: lösliche Fraktion nach Zellyse; P: unlösliche Fraktion nach Zellyse; +: Positivkontrolle (SUMO-AQ mit aminoterminalen (His)<sub>6</sub>-tag).

Auch wenn nur die Bildung einiger *flank*-Proteine nachgewiesen werden konnte, so zeigten die initialen Ergebnisse zur rekombinanten Produktion von *flank*-Proteinen in *E. coli* doch, dass deren heterologe Produktion in einem prokaryotischen Wirtsorganismus prinzipiell möglich war.

### 3.2.2 Herstellung optimierter preCol *flank*-Proteine

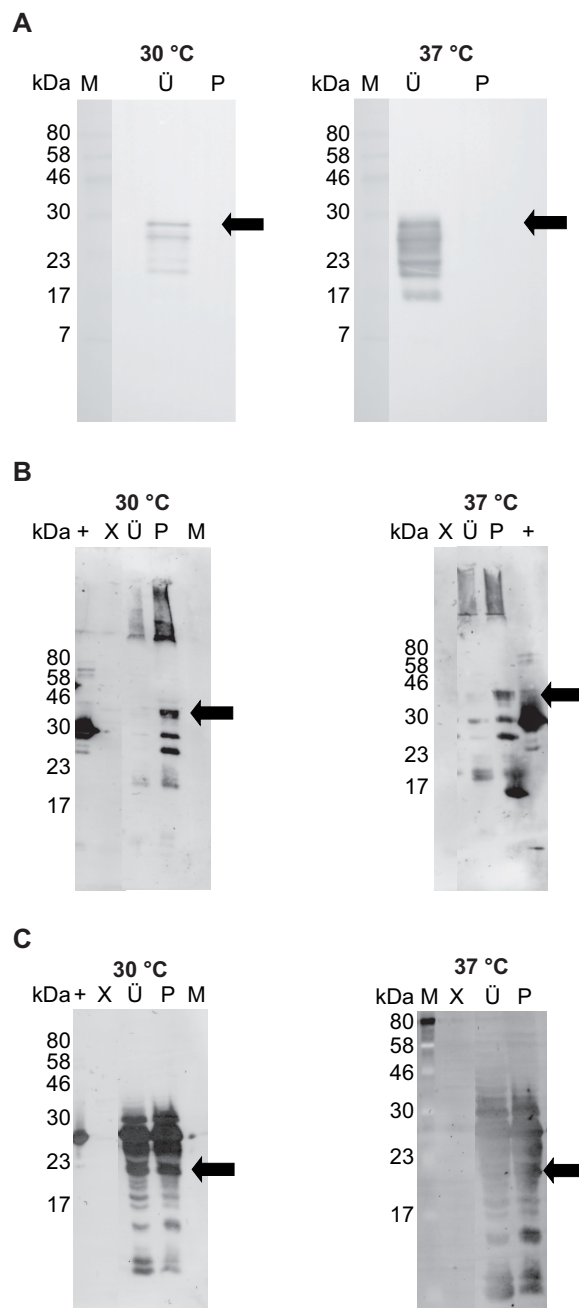
Zusätzlich zu den unveränderten preCol *flank*-Proteinen wurden optimierte *flank*-Proteine hergestellt. Bei diesen handelte es sich um Tyrosin- sowie Tryptophanmutanten. Da die flankierenden Regionen natürlicherweise nur Phenylalanin umfassen, war beabsichtigt, durch das Einbringen der Aminosäuren Tyrosin und Tryptophan mit einem höheren Absorptionskoeffizienten eine Verbesserung der UV-spektroskopischen Eigenschaften für strukturelle sowie funktionelle Analysen zu erreichen. Zudem sollte durch die Verwendung des gefalteten SUMO-tags anstelle des (His)<sub>6</sub>-tags simultan eine Verbesserung der Löslichkeit der rekombinant erzeugten Proteine erreicht werden<sup>[169,196,197]</sup>.

Zur Herstellung der Tyrosinmutanten SUMO-eCT<sub>2Y</sub>, -cwNT<sub>2Y</sub><sup>6</sup> und -cwCT<sub>2Y</sub><sup>6</sup> (Abb. 3-4) konnten die bereits im Expressionsvektor pET28a vorliegenden, der Mfl-1 entstammenden Nukleotidsequenzen (vgl. Kapitel 3.2.1) durch die Wahl entsprechender Primer zunächst so amplifiziert werden, dass an den 3'-Enden in allen Fällen ein Tyrosin- sowie Alanincodon und an den 5'-Enden von cwNT und cwCT ein Tyrosincodon angehängt wurde. Das 5'-Tyrosincodon der eCT<sub>2Y</sub> Sequenz konnte durch primervermittelte Mutagenese des ursprünglich an dieser Stelle vorkommenden Phenylalanincodons erzeugt werden. Die erfolgreiche Ligation der drei Sequenzen in den Expressionsvektor pET28 SUMO (Abb. 2-1), durch welchen der aminoterminal SUMO-*tag* codiert wird, konnte durch Sequenzierung bestätigt werden. Anschließend wurde die heterologe Produktion der optimierten *flank*-Proteine in den *E. coli*-Stämmen BL21-Gold(DE3) und BL21-CodonPlus(DE3)-RIPL bei verschiedenen Temperaturen untersucht.

Die besten Resultate bei der rekombinanten Produktion von SUMO-eCT<sub>2Y</sub>, -cwNT<sub>2Y</sub> und -cwCT<sub>2Y</sub> konnten im Stamm BL21-CodonPlus(DE3)-RIPL erzielt werden (Abb. 3-6). Nach Lyse der Zellen zeigten gegen die am SUMO-*tag* befindliche (His)<sub>6</sub>-Sequenz gerichtete Westernblotanalysen, dass sich das gebildete SUMO-eCT<sub>2Y</sub> ausschließlich in der löslichen Fraktion der Zelllysate befand (Abb. 3-6 A) und dass zudem die Expressionstemperatur keinen nachweisbaren Einfluss auf die Bildung von SUMO-eCT<sub>2Y</sub> hatte. Westernblotanalysen der SUMO-cwNT<sub>2Y</sub> Expressionen zeigten, dass auch hier kein Unterschied zwischen den beiden Expressionstemperaturen erkennbar war. Im Gegensatz zu SUMO-eCT<sub>2Y</sub> lag gebildetes SUMO-cwNT<sub>2Y</sub> jedoch hauptsächlich in der unlöslichen Fraktion der Zellen vor (Abb. 3-6 B). In *E. coli* BL21-CodonPlus(DE3)-RIPL produziertes SUMO-cwCT<sub>2Y</sub> wurde ebenfalls unabhängig von der Expressionstemperatur gebildet und konnte nach Lyse der Zellen sowohl in der löslichen als auch der unlöslichen Fraktion der Zelllysate nachgewiesen werden (Abb. 3-6 C).

---

<sup>6</sup> Die rekombinante Herstellung von SUMO-cwNT<sub>2Y</sub> und -cwCT<sub>2Y</sub> erfolgte durch Martina B. Elsner im Rahmen ihrer durch mich betreuten und an der Universität Bayreuth eingereichten Masterarbeit<sup>[8]</sup>. Entsprechende Daten und Abbildungen sind aus dieser entnommen, modifiziert und zusammengefasst.



**Abb. 3-6. Westernblot zur Analyse der rekombinanten Herstellung von *flank*-Protein-Tyrosinmutanten in *E. coli* BL21-CodonPlus-RIPL. (A) SUMO-eCT<sub>2Y</sub> bei 30 °C (links) und 37 °C (rechts). (B) SUMO-cwNT<sub>2Y</sub> bei 30 °C (links) und 37 °C (rechts). (C) SUMO-cwCT<sub>2Y</sub> bei 30 °C (links) und 37 °C (rechts). Gebildete Fusionsproteine wurden über den aminoterminalen (His)<sub>6</sub>-tag nachgewiesen. Das jeweilige *flank*-Fusionsprotein ist durch einen schwarzen Pfeil markiert M: *NEB ColorPlus Prestained Marker* (7–175 kDa); X: Probe vor Induktion der Genexpression; Ü: lösliche Fraktion der Zelllyse 4 h nach Expressionsstart; P: unlösliche Fraktion der Zelllyse 4 h nach Expressionsstart; +: Positivkontrolle (Ulp1, 27,5 kDa).**

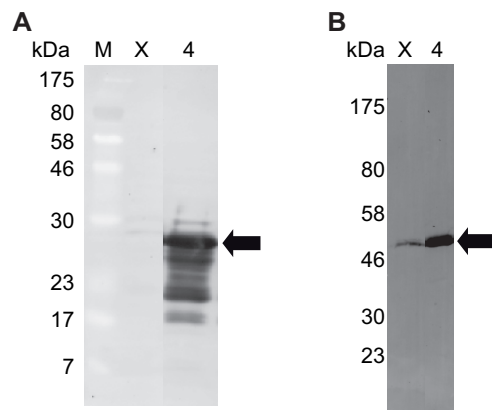
Die Westernblotanalysen zeigten zudem, dass alle drei *flank*-Protein-Tyrosinmutanten bereits während der heterologen Produktion Abbauprozessen unterlagen. Unterhalb der für das jeweilige Fusionsprotein charakteristischen Bande waren weitere Banden, die auf zusätzliche, einen (His)<sub>6</sub>-tag tragende Proteine hindeuteten, zu sehen (Abb. 3-6). Da diese Banden auch schon in während der Expressionskinetiken entnommenen Proben, welche nicht lysiert wurden, auftraten (nicht gezeigt), mussten diese Abbauprozesse bereits intrazellulär stattgefunden haben. Darüber hinaus waren nach Lyse der SUMO-cwNT<sub>2Y</sub> und SUMO-cwCT<sub>2Y</sub> Expressionsproben in den Westernblots auch Signale oberhalb der

Laufweite der jeweiligen *flank*-Proteine sichtbar (Abb. 3-6 B & C). Da auch diese schon vor der Zellyse in Expressionskinetikproben nachweisbar waren (nicht gezeigt), handelte es sich bei ihnen um intrazellulär gebildete, höhermolekulare Aggregate der *flank*-Proteinen bzw. derer Abbauprodukte.

Die zur rekombinanten Herstellung der Tryptophanmutanten SUMO-eCT(F106W), -cwCT(F34W) und -cwCT(F34W)-CBD (Abb. 3-4) verwendeten Nukleotidsequenzen waren an die Codonverwendung von *E. coli* angepasst worden. Die für eCT(F106W) und cwCT(F34W) codierenden artifiziellen Nukleotidsequenzen konnten direkt in den pET28 SUMO-Expressionsvektor ligiert werden, was die Herstellung der Fusionsproteine SUMO-eCT(F106W) und -cwCT(F34W) ermöglichte. Zur Erzeugung des für SUMO-cwCT(F34W)-CBD codierenden Vektors wurde die cwCT(F34W)-Nukleotidsequenz zunächst erfolgreich in den Vektor pTXB1 (Abb. 2-1) ligiert, aus diesem zusammen mit der 3'-CBD-Sequenz durch PCR amplifiziert und die resultierende cwCT(F34W)-CBD Sequenz schließlich in den pET SUMO Expressionsvektor ligiert.

Die besten Ergebnisse hinsichtlich der rekombinanten Produktion wurden für SUMO-eCT(F106W) in *E. coli* BL21-Gold(DE3) und für SUMO-cwCT(F34W)-CBD in *E. coli* BLR(DE3) bei 30 °C erzielt. Die Bildung von SUMO-cwCT(F34W) schlug hingegen in beiden Stämmen fehl. Vier Stunden nach Induktion der Genexpression konnten sowohl SUMO-eCT(F106W) als auch SUMO-cwCT(F34W)-CBD durch Westernblotanalysen nachgewiesen werden (Abb. 3-7). Unterhalb der von SUMO-cwCT(F34W) stammenden Bande waren in den Westernblots wieder zahlreiche weitere Banden detektierbar (Abb. 3-7 A), was, wie auch schon bei der rekombinanten Produktion der Tyrosinmutanten (Abb. 3-6), auf intrazellulären Abbau von SUMO-cwCT(F34W) schließen ließ. Demgegenüber war in der Westernblotanalyse der SUMO-cwCT(F34W)-CBD Expressionsprobe nur eine einzige Bande zu erkennen, was auf ein Ausbleiben intrazellulären Proteinabbaus bei Verwendung zweier terminaler, strukturierter *tags* hindeutete.





**Abb. 3-7. Westernblotanalyse der rekombinanten Herstellung von SUMO-eCT(F106W) und SUMO-cwCT(F34W)-CBD bei 30 °C. (A)** Herstellung von SUMO-eCT(F106W) (25,3 kDa, Pfeil) in *E. coli* BL21-Gold(DE3). **(B)** Herstellung von SUMO-cwCT(F34W)-CBD in *E. coli* BLR(DE3) (48,4 kDa, Pfeil). Die Fusionsproteine wurden über den aminoterminalen (His)<sub>6</sub>-tag nachgewiesen. M: NEB ColorPlus Prestained Marker (7-175 kDa); X: Probe vor Induktion der Genexpression; 4: Probe nach 4 h Expressionszeit.

Nach ihrer Bildung im jeweiligen *E. coli*-Stamm lagen sowohl SUMO-eCT(F106W) wie auch SUMO-cwCT(F34W)-CBD nach Lyse der Bakterienzellen in der löslichen sowie unlöslichen Fraktion vor (nicht gezeigt). Ein Vergleich der Abbildungen 3-6 und 3-7 legt nahe, dass durch Verwendung von an die *E. coli*-Codonverwendung angepasster Nukleotidsequenzen weder eine Steigerung der Menge an heterolog gebildetem Protein noch eine Verringerung der Menge an intrazellulär entstandenen, verkürzten Proteinfragmenten erreicht wurde. Nur in Gegenwart eines zweiten, strukturierten tags (des carboxyterminalen CBD-tags) war die Bildung verkürzter Proteinfragmente nicht mehr zu beobachten (Abb. 3-7 B).

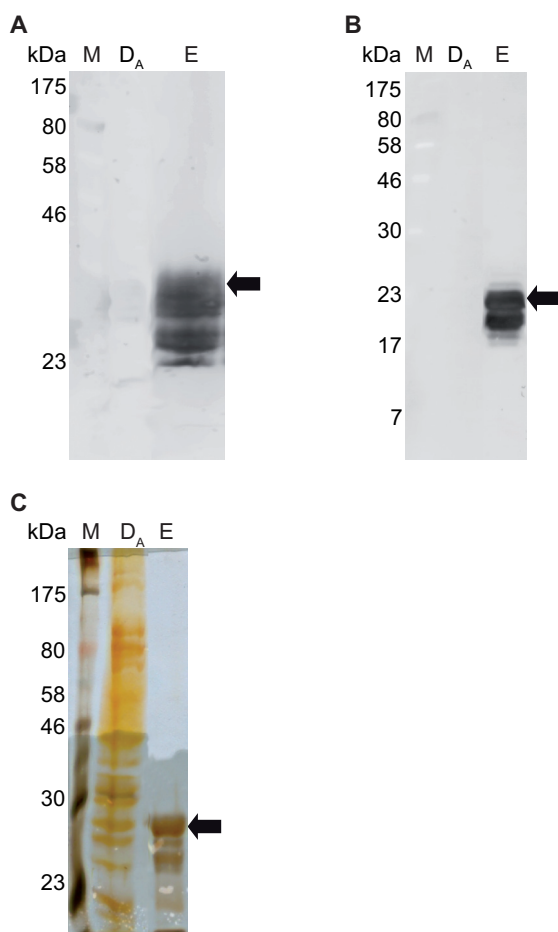
Zusammenfassend konnte gezeigt werden, dass sich die Verwendung von größeren, globulären tags (SUMO und CBD) anstelle der (His)<sub>6</sub>-tags sowohl positiv auf die Löslichkeitseigenschaften der *flank*-Fusionsproteine (vgl. Kapitel 3.2.1) wie auch – im Fall der simultanen Verwendung beider tags – auf eine Degradation des Proteins während der rekombinanten Produktion auswirken kann.

### 3.3 Reinigung rekombinant hergestellter SUMO-*flank*-Fusionsproteine

Da durchgeführte bioinformatische Sekundärstrukturvorhersagen ein wahrscheinliches Auftreten von Sekundärstrukturmotiven in den flankierenden Regionen der *M. galloprovincialis* preCol-Moleküle zeigten (Abb. 3-2), wurden die Reinigung rekombinant hergestellten *flank*-Fusionsproteine bevorzugt unter nicht-denaturierenden Bedingungen aus den löslichen Überständen der Zelllyse vorgenommen.

### 3 Ergebnisse

Die Abtrennung der rekombinanten *flank*-Fusionsproteine von den *E. coli*-eigenen Proteinen erfolgte über Affinitätschromatographie (IMAC, von engl. *immobilized metal-ion affinity chromatography*) über die aminoterminal am SUMO-tag befindliche (His)<sub>6</sub>-Sequenz. Vermittelt durch diese sind die *flank*-Proteine in der Lage, reversibel an eine mit Ni<sup>2+</sup> beladene Sepharosematrix zu binden. Nachdem die *flank*-Fusionsproteine an die Affinitätsmatrix gebunden und nichtbindende *E. coli*-Proteine durch Waschschriffe entfernt worden waren, konnte die Elution der gebundenen Proteine kompetitiv mittels Imidazol erfolgen.



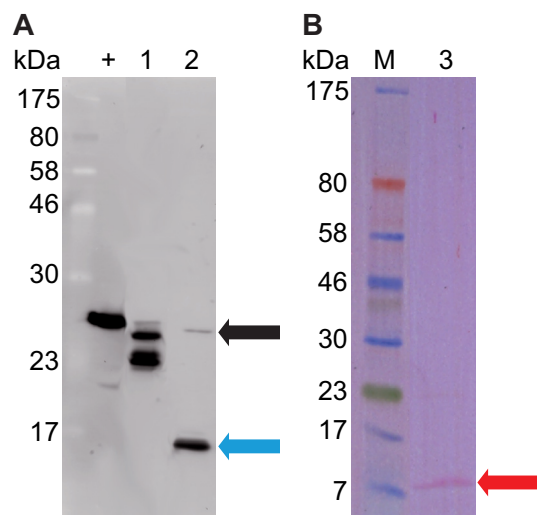
**Abb. 3-8. Analyse der Reinigung der rekombinant hergestellten *flank*-Proteine SUMO-eCT<sub>2Y</sub>, -cwCT<sub>2Y</sub> und -eCT(F106W) mittels IMAC. (A)** Westernblotanalyse der IMAC von SUMO-eCT<sub>2Y</sub>. **(B)** Westernblotanalyse der IMAC von SUMO-cwCT<sub>2Y</sub>. Die Fusionsproteine wurden über den aminoterminalen (His)<sub>6</sub>-tag nachgewiesen. +: Positivkontrolle (Ulp1, 27,5 kDa). **(C)** SDS-PAGE Analyse (15 %iges, silbergefärbtes PAA-Gel) der IMAC von SUMO-eCT(F106W). Das jeweilige *flank*-Fusionsprotein ist durch einen schwarzen Pfeil markiert. M: *NEB ColorPlus Prestained Marker* (7–175 kDa); D<sub>A</sub>: Durchflussfraktion des Säulenauftrags; E: Eluatfraktion des jeweiligen *flank*-Fusionsproteins.

Die Bindung der im Überstand der Zellyse befindlichen *flank*-Fusionsproteine SUMO-eCT<sub>2Y</sub><sup>7</sup>, -cwCT<sub>2Y</sub><sup>7</sup> und -eCT(F106W) an die Affinitätsmatrix verlief in allen Fällen

<sup>7</sup> Die Reinigung von SUMO-cwCT<sub>2Y</sub> erfolgte durch Martina B. Elsner im Rahmen ihrer durch mich betreuten und an der Universität Bayreuth eingereichten Masterarbeit<sup>[8]</sup>. Entsprechende Daten und Abbildungen sind aus dieser entnommen, modifiziert und zusammengefasst.

quantitativ, da die während der Säulenbeladung gesammelten Durchflussfraktionen in den Westernblotanalysen keine Signale zeigten (Abb. 3-8 A & B). Wie zudem aus dem gezeigten, repräsentativen SDS-PAGE Gel der IMAC-Fractionen der SUMO-eCT(F106W) Reinigung (Abb. 3-8 C) ersichtlich ist, gelang durch IMAC eine fast vollständige Abtrennung des jeweiligen *flank*-Fusionsproteins von den in den Zelllysaten enthaltenen *E. coli*-Proteine. Proteinbanden sowie Signale in den Westernblotanalysen unterhalb des jeweiligen Fusionsproteins (Abb. 3-8) legten nahe, dass in allen Eluatfraktionen neben den Volllängenproteinen auch die bereits während der Biosynthese der Proteine entstandenen Fragmente (Abb. 3-6 A & C u. Abb. 3-7 A) vorlagen. Die große Anzahl an fragmentiertem Protein sowie die geringen Größenunterschiede zum entsprechenden *flank*-Fusionsprotein verhinderten eine weitere Reinigung der Eluatfraktionen mittels Gelpermeationschromatographie (GPC).

Dass eine Abspaltung des aminoterminalen SUMO-tags nach Reinigung der SUMO-*flank*-Fusionsproteine nahezu quantitativ möglich war, konnte ferner am Beispiel SUMO-cwCT<sub>2Y</sub> gezeigt werden (Abb. 3-9).



**Abb. 3-9. Analyse der proteolytischen Abspaltung des SUMO-tags von SUMO-cwCT<sub>2Y</sub>.**

**(A)** Westernblotanalyse der SUMO-cwCT<sub>2Y</sub> enthaltenden IMAC-Fractionen vor (1) sowie nach (2) Zugabe der SUMO-Protease Ulp-1. Nach der Spaltung waren nur noch der SUMO-tag sowie die (His)<sub>6</sub>-tag tragende Ulp-1 detektierbar. Schwarzer Pfeil: (His)<sub>6</sub>-Ulp-1; blauer Pfeil: SUMO-tag, +: Positivkontrolle (Ulp1, 27,5 kDa). **(B)** SDS-PAGE Analyse der Durchflussfraktion nach affinitätschromatographischer Abtrennung von (His)<sub>6</sub>-Ulp-1 und SUMO-tag (3). Die Proteine wurden nach elektrophoretischem Transfer auf eine PVDF-Membran mit Ponceau S visualisiert. Roter Pfeil: cwCT<sub>2Y</sub> (7,3 kDa); M: NEB ColorPlus Prestained Marker (7–175 kDa)

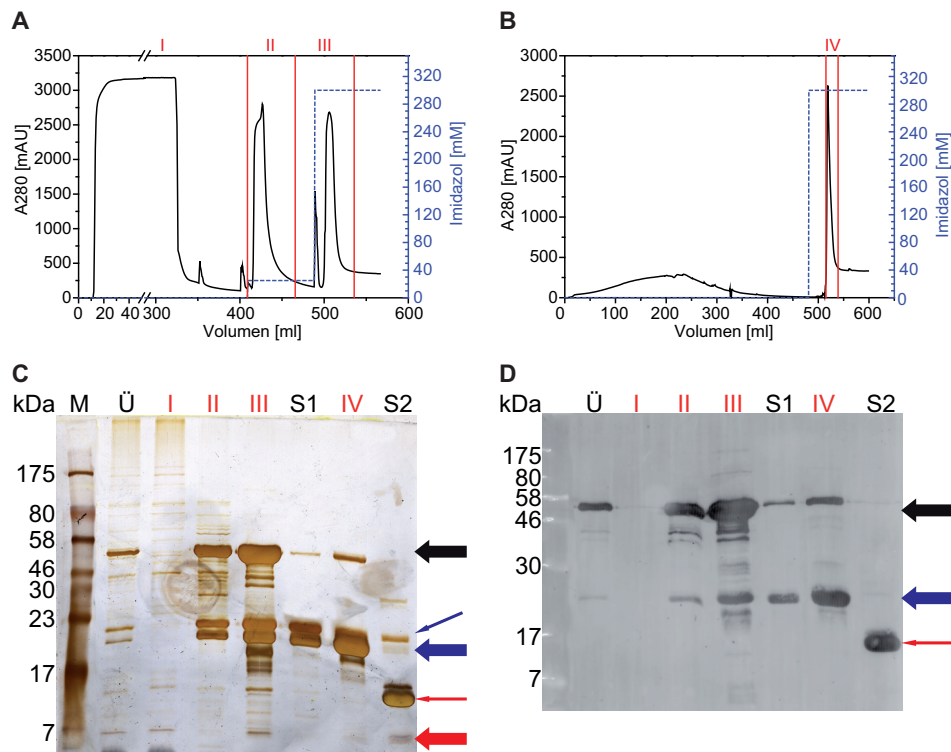
Nach Zugabe der SUMO-Protease Ulp-1 waren in der Westernblotanalyse statt der zahlreichen, von SUMO-cwCT<sub>2Y</sub> sowie dessen Fragmenten stammenden Banden lediglich noch zwei Banden detektierbar, welche dem abgespaltenem SUMO-tag sowie (His)<sub>6</sub>-Ulp-1 entsprachen (Abb. 3-9 A, Spur 2, schwarzer & blauer Pfeil). Der abgespaltene tag wie auch die SUMO-Protease konnten anschließend durch IMAC aus den Ansätzen entfernt

werden und *tag*-freies eCT<sub>2Y</sub> war nach Transfer auf eine PVDF-Membran durch Ponceau S Färbung nachweisbar (Abb. 3-9 B, Spur 3, roter Pfeil). Obwohl neben cwCT<sub>2Y</sub> keine weiteren Proteine in der Probe mehr nachgewiesen werden konnten, wurde dennoch davon ausgegangen, dass immer noch verkürzte Fragmente (analog denen von SUMO-cwCT<sub>2Y</sub> vor der enzymatischen Abspaltung des SUMO-*tags*) in den Proben vorhanden waren. Deshalb waren eindeutige Aussagen zur Reinheit von cwCT<sub>2Y</sub> sowie zu dessen Ausbeute nicht möglich.

#### 3.4 Produktion rekombinanter *flank*-Proteine durch Herstellung und Reinigung doppelt getagter Fusionsproteine

Die sowohl während der rekombinanten Produktion der unveränderten *flank*-Proteine als auch der SUMO-*flank*-Fusionsproteine auftretende Fragmentierung verhinderte die Reinigung des jeweiligen *tag*-freien *flank*-Proteins. Da dieses Phänomen durch die Verwendung eines zweiten, globulären *tags*, der carboxyterminal zur eigentlichen *flank*-Sequenz angebracht wurde, umgangen werden konnte (Abb. 3-7 B), konnte schließlich eine Reinigungsstrategie für *flank*-Proteine am Beispiel von SUMO-cwCT(F34W)-CBD (Abb. 3-4) entwickelt werden.

SUMO-cwCT(F34W)-CBD wurde zunächst rekombinant im *E. coli*-Stamm BLR(DE3) hergestellt. Die Reinigung von cwCT(F34W) erfolgte dann ausgehend von in der löslichen Zellfraktion befindlichem SUMO-cwCT(F34W)-CBD (Abb. 3-10 C & D). Die im Überstand ebenfalls per SDS-PAGE und Westernblot nachweisbaren Proteine SUMO-cwCT(F34W) und der „freie“ CBD-*tag* (Abb. 3-10 C & D) waren dabei methodenbedingte Artefakte, die durch das im SDS-Auftragspuffer enthaltene  $\beta$ -Mercaptoethanol und die von diesem induzierte autokatalytische Abspaltung des Intein-CBD-*tags* verursacht worden waren. Der erste Reinigungsschritt erfolgte durch IMAC. Dadurch war es möglich, das im Überstand der Zelllyse befindliche SUMO-cwCT(F34W)-CBD von einem Großteil der *E. coli*-eigenen Proteine, die zum einen im Durchfluss blieben und zum anderen durch die Waschschrte entfernt werden konnten, abzutrennen (Abb. 3-10 A, C & D). Allerdings führte der zur Erhöhung der Stringenz mit 25 mM Imidazol durchgeführte Waschschrte bereits zu kleineren Verlusten von SUMO-cwCT(F34W)-CBD (Abb. 3-10 A, C & D).

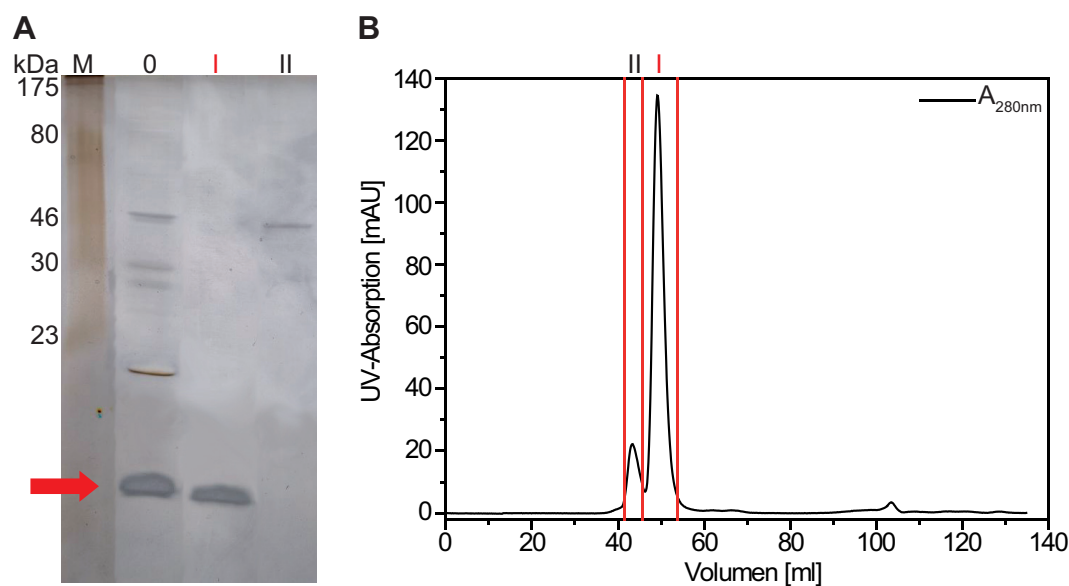


**Abb. 3-10. Chromatogramme, SDS-PAGE- und Westernblotanalyse der zwei initialen Reinigungsschritte während der cwCT(F34W)-Reinigung.** (A) Chromatogramm der IMAC zur Reinigung von (His)<sub>6</sub>-tag Fusionsproteinen aus dem *E. coli* Zelllysat. (B) Chromatogramm der IMAC nach Abspaltung des CBD-tags. (C) SDS-PAGE der ersten beiden Chromatographiestritte der cwCT(F34W)-Reinigung mit anschließender Silberfärbung. (D) Westernblotanalyse der ersten beiden Chromatographiestritte der cwCT(F34W)-Reinigung. Proteine wurden über den aminoterminalen (His)<sub>6</sub>-tag nachgewiesen. Schwarze Pfeile: SUMO-cwCT(F34W)-CBD; blaue Pfeile: SUMO-cwCT(F34W) (dick) u. CBD-tag (dünn); rote Pfeile: cwCT(F34W) (dick) u. SUMO-tag (dünn). M: NEB ColorPlus Prestained Marker (7–175 kDa); Ü: lösliche Fraktion nach Zellyse; S1, S2: Probe nach Abspaltung des CBD- (S1) bzw. SUMO-tags (S2); I-IV: Proben an definierten Stellen der Reinigungsschritte mit entsprechendem Bezug zwischen Chromatogrammen, SDS-PAGE- und Westernblotanalyse.

Nach diesem Reinigungsschritt konnte der carboxyterminale CBD-tag entfernt werden (Abb. 2-2). Das durch DTT induzierte Spleißen der Mini-Intein-Sequenz und die damit verbundene CBD-tag Abspaltung verlief jedoch nicht vollständig quantitativ, so dass selbst nach 48-stündiger Inkubation neben SUMO-cwCT(F34W) und dem abgespaltenem CBD-tag noch Spuren von SUMO-cwCT(F34W)-CBD in den Ansätzen nachweisbar waren (Abb. 3-10 C & D). Der abgespaltene CBD-tag konnte durch eine zweite IMAC abgereichert, jedoch nicht komplett aus dem Ansatz entfernt werden, was auf unspezifische Interaktionen der Chitinbindedomäne mit der Ni<sup>2+</sup>-Affinitätsmatrix hindeutete (Abb. 3-10 B-D). Die anschließend durch Zugabe der SUMO-Protease Ulp1 induzierte Abspaltung des aminoterminalen SUMO-tags verlief im Gegensatz zur Abspaltung des

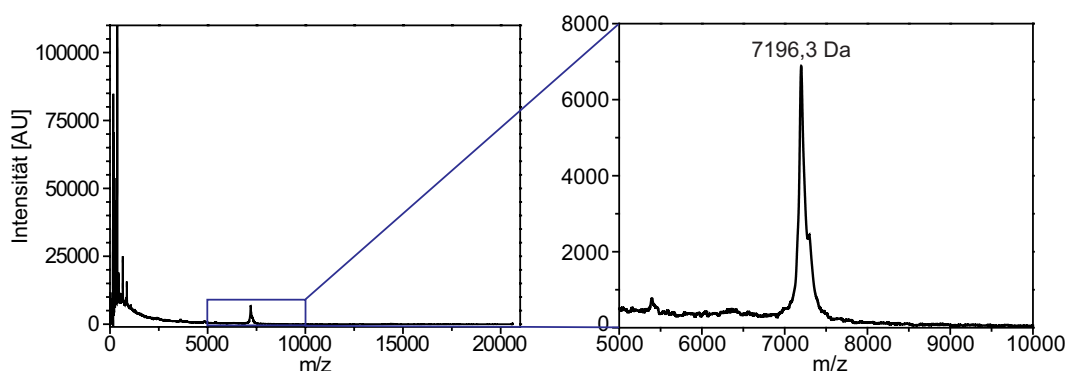
CBD-tags quantitativ (Abb. 3-10 C & D). Während zugesetztes Ulp-1 sowie der abgespaltene SUMO-tag in einem dritten IMAC-Schritt vollständig aus den Ansätzen entfernt werden konnten, verblieb weiterhin ein Teil des abgespaltenen CBD-tags sowie kleinere Mengen cwCT(F34W)-CBD in den Ansätzen (nicht gezeigt).

Aufgrund der Molekulargewichtsunterschiede zwischen dem 35 kDa cwCT(F34W)-CBD bzw. dem 27,9 kDa CBD-tag und dem 7,2 kDa cwCT(F34W) war eine Abtrennung der ersten beiden Proteine von cwCT(F34W) durch eine Gelpermeationsmatrix mit 10 kDa Ausschlussvolumen möglich (Abb. 3-11 A).



**Abb. 3-11. SDS-PAGE Analyse und Chromatogramm der finalen GPC der cwCT(F34W)-Reinigung.** (A) SDS-PAGE Analyse der Probe vor und den Fraktionen nach GPC (Ausschlussvolumen > 10 kDa). M: *NEB Color-Plus Prestained Marker* (7–175 kDa); 0: Proteingemisch vor GPC; I: Fraktion unterhalb der Ausschlussgröße; II: Fraktionen oberhalb der Ausschlussgröße; roter Pfeil: cwCT(F34W). (B) Chromatogramm der GPC mit UV-Absorption bei 280 nm.

Während der GPC eluierten die noch in der Probe befindlichen Proteine in zwei gut voneinander getrennten Fraktionen (Abb. 3-11 B). Eine der beiden Fraktionen enthielt dabei alle Proteine, deren Größe über der Ausschlussgröße der Gelpermeationsmatrix lag, während in der zweiten Fraktion nur noch ein Protein durch SDS-PAGE Analyse nachweisbar war (Abb. 3-11 A, Spur I). Aus der SDS-PAGE konnte geschlossen werden, dass es sich bei dem mit dem zweiten UV-Peak verbundenen, eluierten Protein um cwCT(F34W) hoher Reinheit handelte (Abb. 3-11 A, roter Pfeil).



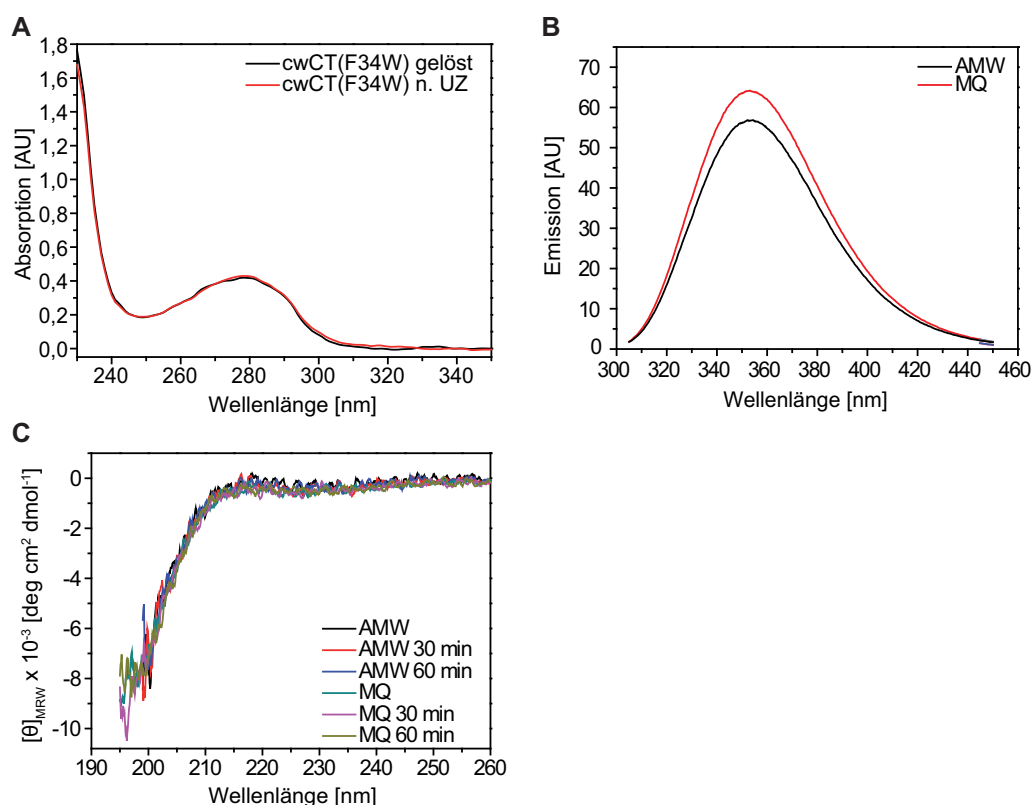
**Abb. 3-12. MALDI-TOF Massenspektrum von gereinigtem cwCT(F34W).** Durch MALDI-TOF konnte das gereinigte Protein als cwCT(F34W) identifiziert sowie dessen Reinheit bestätigt werden.

Zur endgültigen Bestätigung, ob es sich bei dem gereinigten Protein tatsächlich um cwCT(F34W) handelte, wurde die exakte Masse des in der GPC-Fraktion I (Abb. 3-11) enthaltenen Proteins mittels MALDI-TOF Massenspektrometrie bestimmt. Das Massenspektrum war durch einen einzelnen Peak geringer Intensität bei einem Masse-Ladungs-Verhältnis von 7196,3 Da und einer kleinen Schulter bei 7231,3 Da charakterisiert (Abb. 3-12, rechte Seite). Der Peak entspricht dabei der Masse von cwCT(F34W) (7178,5 Da), an welches zusätzlich ein aus dem Puffer stammendes Ammoniumkation (18 Da) gebunden war. Die Schulter des Peaks entspricht cwCT(F34W)-Molekülen, die höchstwahrscheinlich mit drei Ammoniumkationen (54 Da) oder aber einem Ammonium- sowie einem Kaliumkation (57 Da) assoziiert waren. Somit konnte durch MALDI-TOF Massenspektrometrie bestätigt werden, dass es sich bei dem in der GPC-Fraktion, die unterhalb des Ausschlussvolumens der Gelpermeationsmatrix eluierte, enthaltenen Protein tatsächlich um cwCT(F34W) handelte. Das Fehlen weiterer Peaks belegte zudem (Abb. 3-12, linke Seite), dass das isolierte cwCT(F34W) eine hohe Reinheit besaß.

### 3.5 Struktur und Stabilität des *flank*-Proteins cwCT(F34W)

Das nach Reinigung lyophilisierte cwCT(F34W) konnte bei Raumtemperatur in gepufferter wässriger Lösung (40 mM Tris/HCl, pH 7,6), in reinem Milli-Q Wasser (MQ) und in artifiziellem Meerwasser (AMW) gelöst werden. Der qualitative Verlauf des cwCT(F34W) UV-Spektrums des in Tris/HCl-Puffer gelösten cwCT(F34W) mit einer Nulllinie > 320 nm

sowie einem steilen Anstieg der Absorption erst unterhalb von 245 nm deutete darauf hin, dass keine durch Aggregate verursachten Streueffekte auftraten (Abb. 3-13 A). Demnach hatte sich das in den Puffer eingewogene cwCT(F34W) bei Raumtemperatur vollständig gelöst.

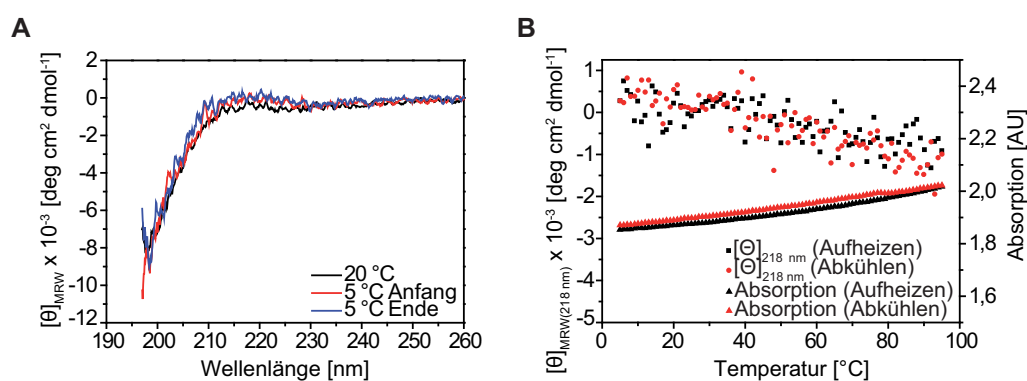


**Abb. 3-13. Löslichkeit und strukturelle Charakterisierung von cwCT(F34W) in wässrigen Lösungen. (A)** UV-Spektren von frisch in 40 mM Tris/HCl, pH 7,6 gelöstem cwCT(F34W) sowie cwCT(F34W) UV-Spektren nach einstündiger Inkubation bei RT mit anschließender Ultrazentrifugation (UZ). **(B)** Tryptophan-Fluoreszenzemissionsspektren von in Milli-Q Wasser (MQ) sowie artifiziellem Meerwasser (AMW) gelöstem cwCT(F34W). **(C)** Fern-UV CD-Spektren von frisch in MQ und AMW bei Raumtemperatur gelöstem cwCT(F34W) sowie nach 30 min und 60 min.

Das gleiche Lösungsverhalten war auch in MQ und AMW nachweisbar. Das sofort nach dem Lösen von cwCT(F34W) aufgenommene UV-Spektrum war darüber hinaus mit dem nach einer Stunde Inkubation und anschließender Ultrazentrifugation aufgenommenen Spektrum (Abb. 3-13 A) identisch. Dies zeigte, dass es innerhalb einer Stunde bei Raumtemperatur nicht zu (UV-spektroskopisch nachweisbarer) Aggregation gekommen war. In MQ und AMW zeigten die Tryptophan-Fluoreszenzemissionsspektren des gelösten cwCT(F34W) ein Emissionsmaximum bei  $\sim 354$  nm (Abb. 3-13 B). Dies bedeutete, dass



das in cwCT(F34W) enthaltene Tryptophan bezüglich seiner wässrigen Umgebung vollständig exponiert war<sup>[189]</sup>. Die zwischen den in MQ und ASW aufgenommenen cwCT(F34W)-Fluoreszenzspektren bestehenden Intensitätsunterschiede waren lösungsmittelbedingt und ließen nicht darauf schließen, dass sich das in cwCT(F34W) enthaltene Tryptophan in MQ bzw. AMW in Umgebungen unterschiedlicher Polarität aufhielt (Abb. 3-13 B). CD-Spektren, die zur qualitativen Bestimmung der Sekundärstruktur von in MQ und AMW gelöstem cwCT(F34W) aufgenommen wurden, waren durch stark negative Werte der molaren Elliptizität bei  $< 200$  nm sowie dem Fehlen von Absorptionsbanden im Bereich oberhalb von 210 nm gekennzeichnet (Abb. 3-13 C). Dies ist typisch für intrinsisch entfaltete Proteine (IDP, von engl. *intrinsically disordered protein*) mit einem gewissen Anteil  $3_1$ -helikaler Strukturen, wie z.B. Polyglycin-II-Helices<sup>[184,185,198]</sup>. Nach Inkubation bei Raumtemperatur aufgenommene Fern-UV CD-Spektren des cwCT(F34W) glichen denen des frisch gelösten Proteins (Abb. 3-13 C). Dies war ein weiterer Beleg für die hohe Stabilität von monomerem cwCT(F34W) in wässrigen Lösungen.



**Abb. 3-14. Thermische Stabilität von cwCT(F34W) in wässriger Lösung.** (A) cwCT(F34W) CD-Spektren vor und nach einem mit 1 °C/min durchgeführten Temperaturübergang. (B) Darstellung der molaren Elliptizität bei 218 nm sowie der zugehörigen UV-Absorption in Abhängigkeit der Temperatur.

Das Aufheizen von in MQ gelöstem cwCT(F34W) mit 1 °C / min auf 95 °C und anschließendes Abkühlen mit derselben Geschwindigkeit führte zu keiner CD-spektroskopisch verfolgbaren Sekundärstrukturänderung. Die vor und nach dem Temperaturübergang aufgenommenen CD-Spektren bei 5 °C glichen einander und entsprachen zudem auch einem zuvor bei Raumtemperatur aufgenommenen Spektrum (Abb. 3-14 A). Die während des Heizvorgangs in 1 °C/min-Schritten detektierte molare Elliptizität bei 218 nm

wies nur geringe, temperaturbedingte Veränderungen auf (Abb. 3-14 B). Dies zeigte, dass cwCT(F34W) in MQ entweder eine vollständig ungeordnete Konformation besaß oder aber eventuell vorhandene  $3_1$ -helikale Konformationen bei thermischer Denaturierung nicht in  $\beta$ -faltblattreiche Aggregate überführt wurden. Dieses Ergebnis wurde zudem durch ähnliche UV-Absorptionswerte vor und nach dem Temperaturübergang untermauert (Abb. 3-14 B).

### 3.6 Strukturelle Interaktion von cwCT(F34W) mit Mizellen und kleinen unimellaren Lipidvesikeln (SUV)

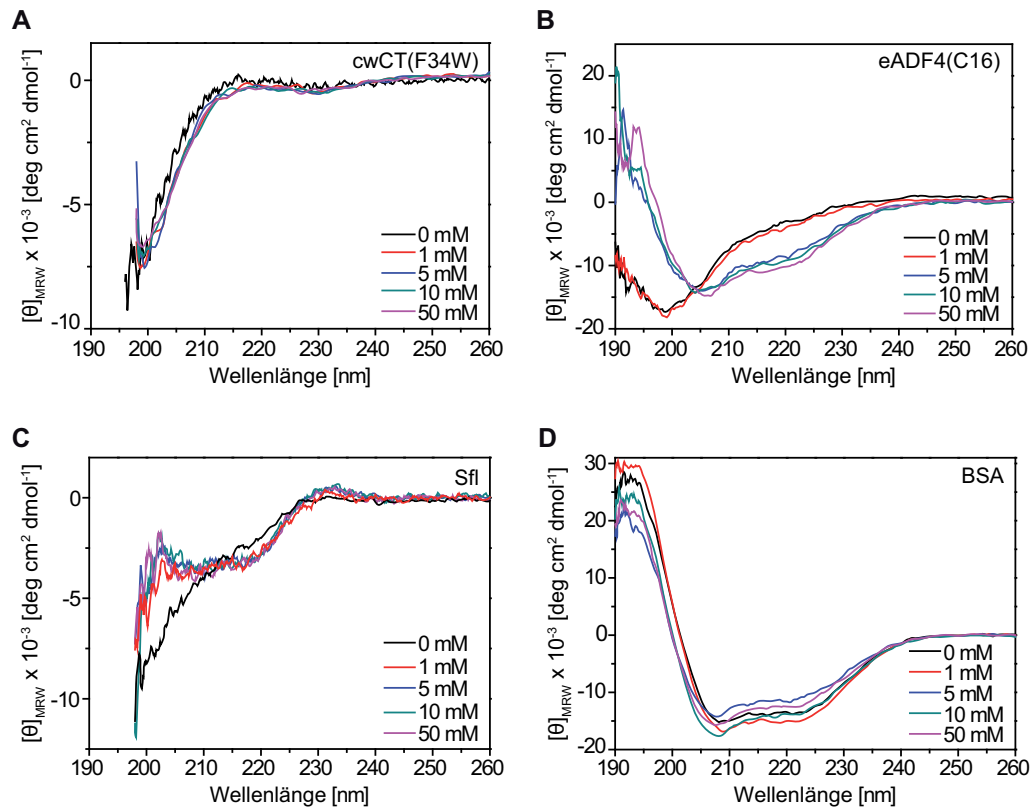
IDPs stellen eine wichtige Proteinklasse mit einer stark wachsenden Anzahl neu entdeckter Vertreter dar<sup>[199,200]</sup>. IDPs selbst können oft mehr als eine einzelne Funktion erfüllen, da sie in der Lage sind, mehrere funktionelle Konformationen einzunehmen<sup>[201-203]</sup>. Jede der möglichen Konformationen ist dabei das Ergebnis spezifischer Interaktionen mit anderen Molekülen, wie Proteinen und Lipiden, oder mit komplexen Oberflächen. Da rekombinant hergestelltes cwCT(F34W) in wässrigen Lösungen eine intrinsisch entfaltete Konformation einnahm (Abb. 3-13 C), wurde untersucht, ob es bei Interaktion mit anderen Molekülen eine für IDPs typische Änderungen seiner Konformation zeigt. Als potentielle Interaktionspartner wurden Lipidmoleküle ausgewählt, da preCol-Moleküle *in vivo* sowohl während ihrer Sekretion in membranumschlossenen sekretorischen *Granulae* (SG) als auch im reifen Byssusfaden mit Lipiden in Kontakt kommen könnten<sup>[31,37,43,123,204,205]</sup>.

#### 3.6.1 Interaktion von cwCT(F34W) mit SDS-Mizellen

Neben dem klassischen Einsatzgebiet von SDS als anionischem Amphiphil und Proteindenaturierungsmittel finden seine Mizellen auch als sehr vereinfachtes Lipidvesikelmodell Verwendung. Deswegen wurde zur Untersuchung möglicher Interaktionen von cwCT(F34W) mit Lipidstrukturen zunächst auf dieses vereinfachte Modellsystem zurückgegriffen. SDS bildet oberhalb seiner kritischen Mizellbildungskonzentration (CMC) Mizellen. Die CMC ist von verschiedenen Faktoren wie Ionenstärke der Lösung, pH-Wert, Temperatur und der Gegenwart anderer Moleküle abhängig und liegt je nach Bestim-

mungsmethode zwischen 0,14–8 mM. Zur Untersuchung potentieller Interaktionen von cwCT(F34W) und SDS wurden deswegen verschiedene SDS-Konzentrationen eingesetzt und die Lösungen zusammen mit löslichem cwCT(F34W) inkubiert. Als Kontrollproteine wurden ein im nativen Zustand gefaltetes Protein (BSA, bovines Serumalbumin) sowie die artifiziellen Spinnenseidenproteine eADF4(C16) und Sfl verwendet, da sich letztere beide, wie cwCT(F34W) auch, von extrakorporalen repetitiven Strukturproteinen ableiten, glycinreiche Konsensusmotive umfassen und in wässrigen Lösungen ebenfalls intrinsisch entfaltet sind<sup>[148,149]</sup>.

Direkt nach dem Mischen wurden CD-Spektren aufgenommen, um Auswirkungen der SDS-Mizellen auf die Sekundärstruktur der untersuchten Proteine nachverfolgen zu können (Abb. 3-15). SDS zeigte im untersuchten Konzentrationsbereich von 1–50 mM keinen strukturellen Einfluss auf cwCT(F34W). Die CD-Spektren, die in Gegenwart von SDS aufgenommen wurden, waren wie das Spektrum von cwCT(F34W) in gepufferter wässriger Lösung durch eine negative molare Elliptizität < 200 nm sowie ein schwaches Maximum bei ~ 220 nm – beides typisch für intrinsisch entfaltete Proteine – gekennzeichnet (Abb. 3-15 A). Im Gegensatz dazu zeigten die als Kontrollen verwendeten, artifiziellen Spinnenseidenproteine eADF4(C16) und Sfl in Gegenwart von SDS Veränderungen ihrer Sekundärstrukturen. Die in Abwesenheit und in Gegenwart von 1 mM SDS aufgenommenen CD-Spektren von eADF4(C16) belegten durch das deutliche Minimum bei < 200 nm, dass das Protein noch seine anfängliche intrinsisch entfaltete Konformation besaß (Abb. 3-15 B). Bei SDS-Konzentrationen  $\geq 5$  mM konnte jedoch eine Änderung der Sekundärstruktur beobachtet werden. Das Minimum bei 208 nm, ein weiteres schwach ausgeprägtes bei ~ 222 nm sowie das Maximum im Bereich 191–193 nm waren typisch für das Vorhandensein  $\alpha$ -helikaler Strukturanteile<sup>[184,185]</sup>. Zudem war zwischen den IDP-typischen und  $\alpha$ -helikal geprägten CD-Spektren ein gemeinsamer Schnittpunkt (isosbestischer Punkt) bei ~ 205 nm vorhanden (Abb. 3-15 A), was zeigte, dass die beiden Konformationszustände direkt ineinander übergangen und die Sekundärstrukturänderung nicht über Konformationszwischenstufen verlief.



**Abb. 3-15. Einfluss von SDS-Mizellen auf die Sekundärstrukturen verschiedener Proteine.** Fern-UV CD-Spektren zur Untersuchung des Einflusses von SDS-Mizellen auf die Sekundärstruktur von cwCT(F34W) (A), eADF4(C16) (B), Sfl (C) sowie BSA (D).

Für das artifizielle Flagelliform-Spinnenseidenprotein Sfl konnte anhand des Auftretens von Minima bei 208 nm und 222 nm ebenfalls ein SDS-bedingter Sekundärstrukturübergang von einer intrinsisch entfalteten zu einer  $\alpha$ -helikal dominierten Konformation beobachtet werden<sup>[184,185]</sup> (Abb. 3-15 C). Der Strukturübergang trat bei Sfl, im Gegensatz zu eADF4(C16), jedoch schon bei SDS-Konzentrationen < 1 mM auf. Unterschiedliche Intensitäten der beiden Minima zwischen den eADF4(C16) und Sfl CD-Spektren und höhere molare Elliptizitäten von eADF4(C16) unterhalb 208 nm waren Indizien dafür, dass die in beiden Proteinen entstandenen  $\alpha$ -helikalen Strukturanteile variierten (Abb. 3-15 C). Die bei beiden artifiziellen Spinnenseidenproteinen beobachtete Strukturänderung fand zudem sehr schnell statt, da alle Spektren sofort nach Mischen der jeweiligen Proteinlösung mit den SDS-Mizellen aufgenommen wurden. Die Konformation von BSA zeigte wiederum keine SDS-bedingten Änderungen und die ursprüngliche,  $\alpha$ -helikale Konformation blieb in Gegenwart der SDS-Mizellen bestehen (Abb. 3-15 D).

Unter den getesteten Strukturproteinen (bzw. Strukturproteinanaloge) war cwCT(F34W) somit das einzige Protein, dessen Konformation von den als stark simplifiziertes Lipid-vesikelmodell verwendeten SDS-Mizellen nicht beeinflusst wurde.

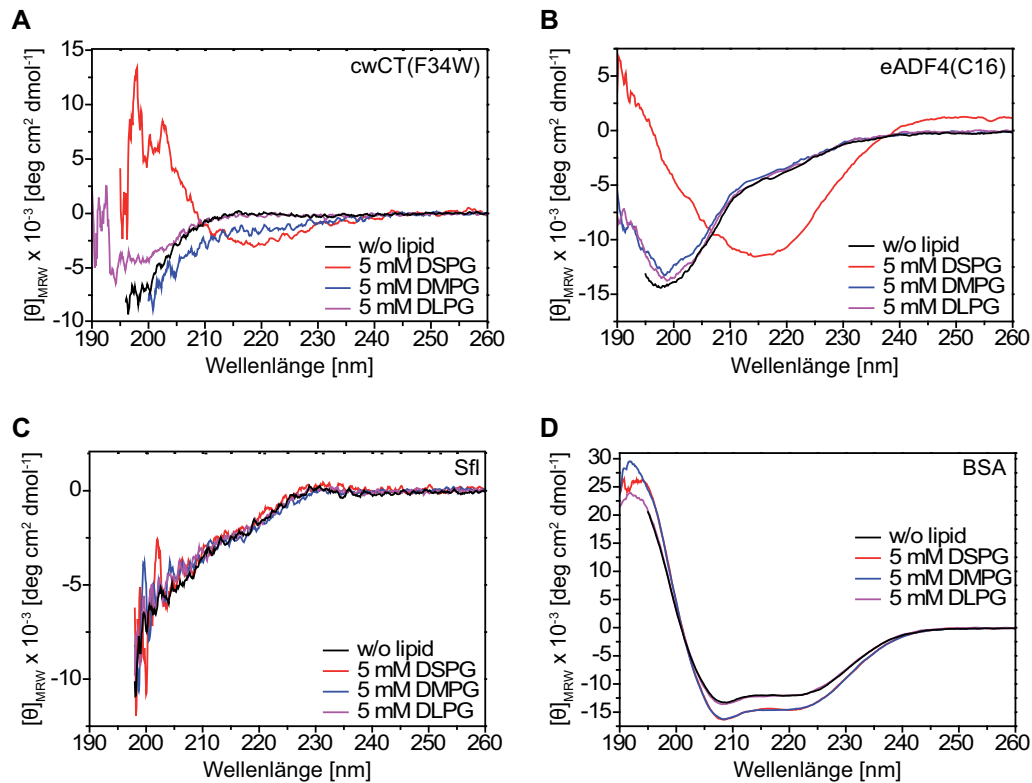
### 3.6.2 Interaktion von cwCT(F34W) mit kleinen unilamellaren Lipidvesikeln (SUV)

Nach der Untersuchung des potentiellen Einflusses eines simplifizierten Vesikelmodells auf die Konformation von cwCT(F34W) wurde in einem nächsten Schritt untersucht, ob cwCT(F34W) mit komplexeren Lipidmembransystemen interagiert. Dazu wurden aus den Phospholipiden DLPG, DMPG und DSPG SUV-Emulsionen in gepufferten wässrigen Lösungen hergestellt. Da SUV aus einer einzelnen Lipiddoppelmembran bestehen, werden sie häufig als Modellsystem zur Untersuchung der Interaktionen von Molekülen mit (biologischen) Lipidmembranen herangezogen.

Die Gegenwart von DLPG-SUV, die aus 5 mM monomerem DLPG erzeugt wurden, hatte weder auf die Sekundärstruktur von cwCT(F34W) noch auf die Sekundärstrukturen der erneut als Kontrollen verwendeten Proteine eADF4(C16), Sfl und BSA einen durch CD-Spektroskopie nachweisbaren Einfluss. In Abwesenheit von DLPG-SUV besaßen sowohl cwCT(F34W) als auch die beiden artifiziiellen Spinnenseidenproteine eine intrinsisch entfaltete Konformation, wie die stark negative molare Elliptizität bei  $< 200$  nm belegte<sup>[184,185]</sup> (Abb. 3-16 A-C). Das CD-Spektrum von BSA blieb in Anwesenheit der DLPG-SUV dem in gepufferter, wässriger Lösung und die ursprüngliche Konformation wurde durch DLPG-SUV genauso wenig beeinflusst wie die von cwCT(F34W) oder der beiden Spinnenseidenproteine (Abb. 3-16 A-D).

DMPG-SUV hatten wie DLPG-SUV keinen Einfluss auf die Konformation der Kontrollproteine eADF4(C16), Sfl und BSA, da in Abwesenheit der SUV aufgenommenen Spektren die gleichen qualitativen Verläufe zeigten wie die jeweiligen, in wässriger gepufferter Lösung aufgenommenen (Abb. 3-16 B-D). cwCT(F34W) hingegen zeigte in Anwesenheit der DMPG-SUV Unterschiede im CD-Spektrum, verglichen mit dem in Abwesenheit der SUV aufgenommenen. Das Auftreten eines lokalen Minimums zwischen 215 nm und 220 nm in Kombination mit immer noch negativen molaren Elliptizitäten im Wellenlängenbereich  $< 200$  nm deutete darauf hin, dass es durch die SUV zur Induktion von

$\beta$ -Faltblattkonformationen, die nun neben der immer noch dominierenden, intrinsisch entfalteten Konformation vorlagen, gekommen sein könnte<sup>[184,185]</sup> (Abb. 3-16 A).

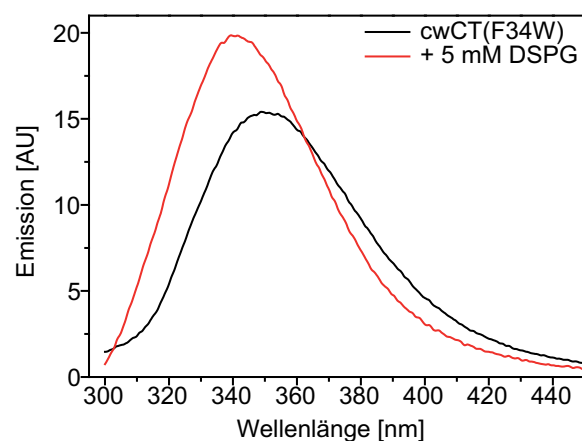


**Abb. 3-16. Fern-UV CD-Spektren verschiedener Proteine in Gegenwart von SUV verschiedener Phospholipide.** Charakterisierung der Sekundärstruktur von cwCT(F34W) (A), eADF4(C16) (B), Sfl (C) und BSA (D) in Abwesenheit sowie in Gegenwart von DLPG-, DMPG- und DSPG-SUV.

Wurden DSPG-SUV verwendet, waren noch weitreichendere konformationelle Änderungen in cwCT(F34W) detektierbar. Ein deutlich stärker ausgeprägtes Minimum bei  $\sim 218$  nm sowie nun positive molare Elliptizität im Bereich 195–200 nm waren klare Indizien, dass der Anteil an induzierter  $\beta$ -Faltblattstruktur, verglichen mit der durch DMPG-SUV verursachten, deutlich höher war<sup>[184,185]</sup> (Abb. 3-16 A). Die Gegenwart von DSPG-SUV führte auch bei eADF4(C16) zu Änderungen der Sekundärstruktur. Die entsprechenden CD-Spektren ließen ebenfalls auf die Entstehung von  $\beta$ -Faltblattstruktur schließen (Abb. 3-16 B). Die CD-Spektren von Sfl sowie BSA wiesen hingegen auch in Gegenwart der DSPG-SUV keine qualitativen Änderungen auf und zeigten, dass die DSPG-SUV keinen Einfluss auf die Konformation beider Proteine hatten (Abb. 3-16 C & D).

Generell mussten alle durch SUV induzierten Konformationsänderungen unmittelbar und schnell erfolgt sein, da die CD-Spektren sofort nach in Kontakt bringen des jeweiligen Proteins mit den SUV aufgenommen wurden.

In Gegenwart von DSPG-SUV gemessene Fluoreszenzemissionsspektren des in cwCT(F34W) enthaltenen Tryptophanrests zeigten, dass das durch die SUV induzierte Auftreten einer  $\beta$ -Faltblattkonformation (Abb. 3-16 A) auch mit Änderungen der chemischen Umgebung des Tryptophans einherging. Das Fluoreszenzemissionsspektrum in wässrigen, gepufferten Lösungen, in denen cwCT(F34W) eine intrinsisch entfaltete Struktur aufwies (Abb. 3-13 C), war durch ein Emissionsmaximum bei  $\sim 350$  nm charakterisiert (Abb. 3-17), was belegte, dass der Tryptophanrest in cwCT(F34W) gegenüber dem polaren Lösungsmittel vollständig exponiert war.



**Abb. 3-17. Tryptophan-Fluoreszenzemissionsspektren von cwCT(F34W).** Die Anregung der Fluoreszenz des in cwCT(F34W) befindlichen Tryptophans erfolgte in Abwesenheit sowie Anwesenheit von DSPG-SUV bei 280 nm.

In Gegenwart von aus 5 mM DSPG hergestellten SUV waren sowohl eine Blauverschiebung (hypsochromer Effekt) der Position des Maximums hin zu einer Wellenlänge von  $\sim 341$  nm als auch eine Zunahme der Fluoreszenzintensität nachweisbar (Abb. 3-17). Beides unterstrich, dass sich das in der cwCT(F34W)-Primärstruktur enthaltene Tryptophan in Gegenwart von DSPG-SUV in einer weniger polaren Umgebung befand und nicht mehr vollständig zum wässrigen Lösungsmittel exponiert vorlag<sup>[189]</sup>.

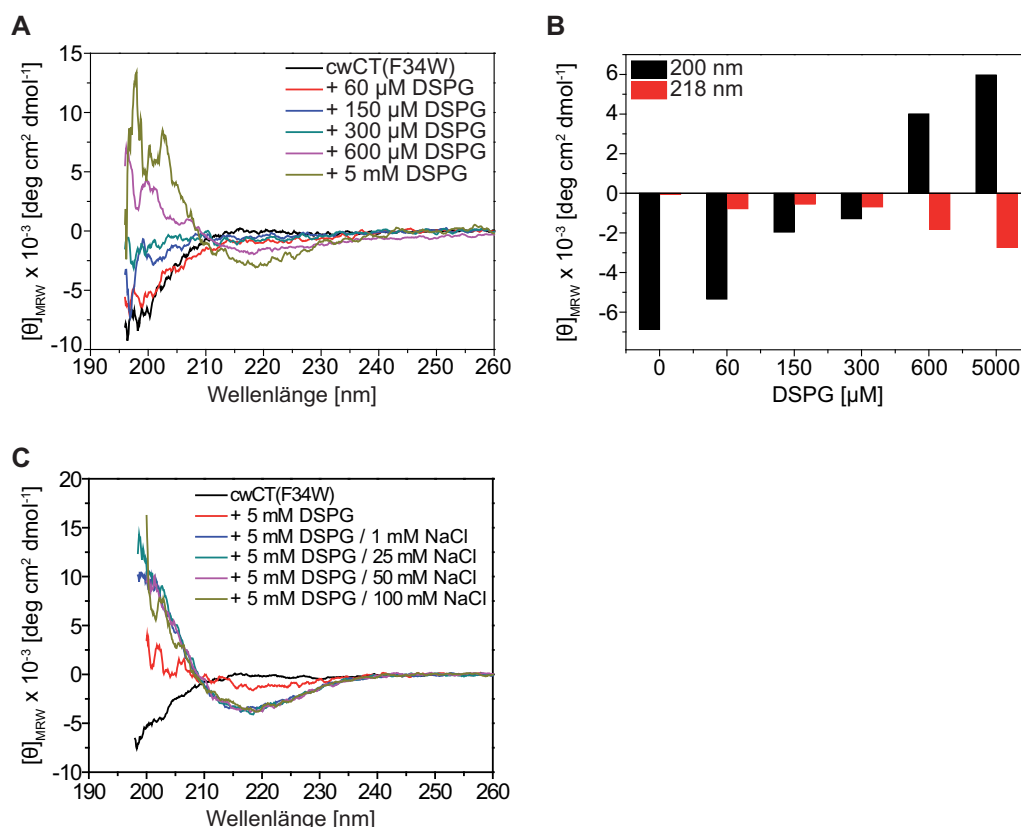
#### **3.6.3 Abhängigkeit der DSPG-SUV-induzierten cwCT(F34W)-Sekundärstrukturänderung von verschiedenen Parametern**

Zur näheren Charakterisierung des strukturinduzierenden Effekts von DSPG-SUV auf cwCT(F34W) wurde untersucht, welchen Einfluss das Verhältnis von DSPG-SUV-Konzentration zu Proteinkonzentration sowie die Ionenstärke auf das Ausmaß der Sekundärstrukturänderung hatte und ob die induzierte  $\beta$ -Faltblattstruktur transienter oder permanenter Natur war.

Zur Analyse der Auswirkungen unterschiedlicher DSPG-SUV-zu-Proteinmolekül-Verhältnisse auf die Sekundärstrukturinduktion in cwCT(F34W) wurden CD-Spektren des Proteins in Abwesenheit sowie in Gegenwart unterschiedlicher DSPG-SUV-Konzentrationen aufgenommen. Um dies zu gewährleisten wurden unterschiedliche DSPG-Ausgangskonzentrationen (oberhalb der CMC) zur Herstellung der SUV verwendet, da die DSPG-Monomerkonzentration bei konstanten Versuchsbedingungen direkt mit der Anzahl entstehender SUV korreliert. Der qualitative Verlauf der CD-Spektren (Abb. 3-18 A) sowie eine Gegenüberstellung der DSPG-Monomerkonzentrationen und der molaren Elliptizitäten bei 200 nm und 218 nm (Abb. 3-18 B), welche signifikante Änderungen beim Übergang von einer intrinsisch entfalteten zu einer  $\beta$ -Faltblattkonformation zeigen, offenbarten bei gleichbleibender cwCT(F34W)-Konzentration einen Zusammenhang zwischen DSPG-SUV-Konzentration und Ausmaß an induzierter  $\beta$ -Faltblattstruktur.

Die intrinsisch entfaltete Konformation des cwCT(F34W) wurde in Gegenwart von DSPG-SUV aus 60  $\mu$ M DSPG noch nicht beeinflusst, das CD-Spektrum zeigte weiterhin einen für intrinsisch entfaltete Proteine typischen Verlauf und die molaren Elliptizitäten bei 200 nm und 218 nm wiesen, bezogen auf jene in Abwesenheit der DSPG-SUV detektierten, nur geringe Unterschiede auf (Abb. 3-18 A & B). Im Gegensatz dazu induzierten die aus 600  $\mu$ M DSPG hergestellten SUV markante Konformationsänderungen und das Minimum bei  $\sim 218$  nm sowie positive molare Elliptizitäten zwischen 195 nm und 200 nm deuteten auf die Ausbildung von  $\beta$ -Faltblattstrukturen in cwCT(F34W) hin (Abb. 3-18 A & B).





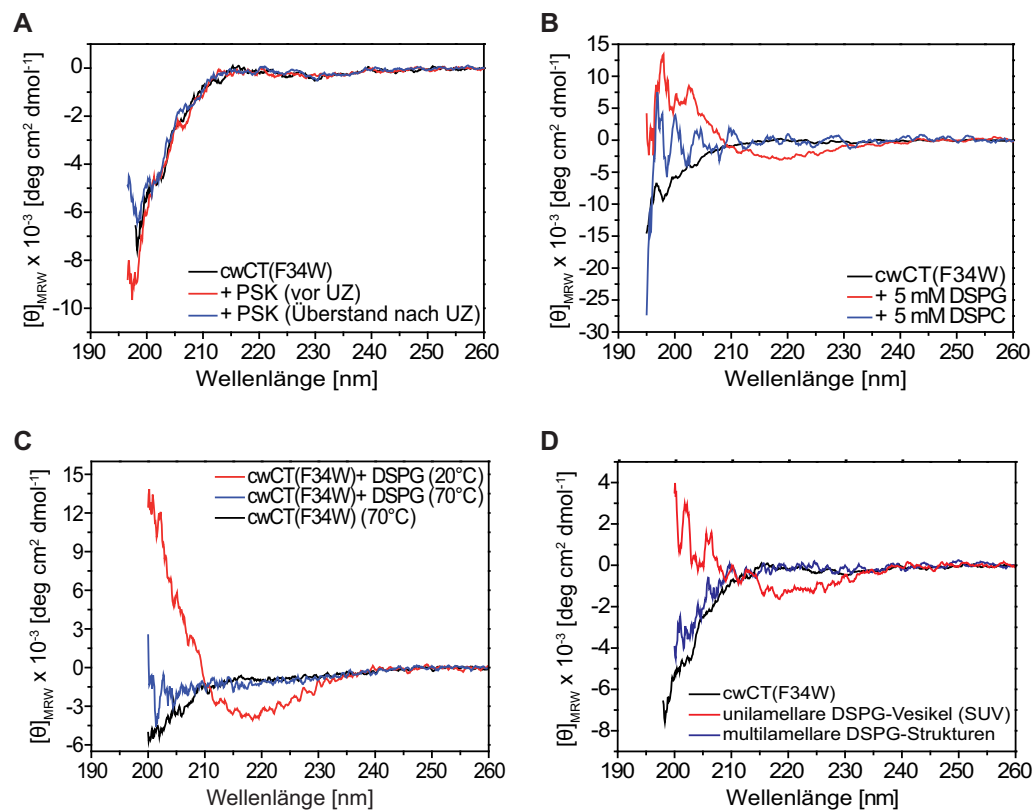
**Abb. 3-18. Fern-UV-CD-Spektren zur Analyse des Einflusses von DSPG-SUV-Konzentration, Ionenstärke und Temperatur auf die Induktion der  $\beta$ -Faltblattstruktur in cwCT(F34W).** (A) Spektren von cwCT(F34W) in Gegenwart unterschiedlicher DSPG-SUV-Konzentrationen. Die Angaben beziehen sich dabei auf die DSPG-Monomerkonzentration. (B) Änderung der molaren Elliptizität bei 200 nm und 218 nm in Abhängigkeit der verwendeten DSPG-Monomerkonzentration. (C) cwCT(F34W)-Spektren in Gegenwart von DSPG-SUV bei verschiedenen NaCl-Konzentrationen.

Der Vergleich der molaren Elliptizitäten bei 218 nm zeigte, dass der  $\beta$ -Faltblattanteil noch unter dem lag, der bei Verwendung einer DSPG-Ausgangskonzentration von 5 mM induziert wurde (Abb. 3-18 A & B). Die bei DSPG-Ausgangskonzentrationen von 150  $\mu\text{M}$  und 300  $\mu\text{M}$  aufgenommenen cwCT(F34W)-CD-Spektren ähnelten weder denen der intrinsisch entfalteten noch denen der  $\beta$ -faltblattreichen Konformation von cwCT(F34W) (Abb. 3-18 A) und waren v. a. durch eine Zunahme der molaren Elliptizitäten im Bereich um 200 nm gekennzeichnet (Abb. 3-18 B).

Veränderungen der Ionenstärke durch Variation der NaCl-Konzentration hatten zur Folge, dass der durch DSPG-SUV ausgelöste  $\beta$ -faltblattstrukturinduzierende Effekt verglichen mit den NaCl-freien Proben apparent weiter verstärkt wurde. Dies war in den cwCT(F34W) CD-Spektren daran erkennbar, dass in Gegenwart von NaCl das für

$\beta$ -Faltblattkonformationen charakteristische Minimum bei 218 nm ausgeprägt war und die molare Elliptizität bei  $< 200$  nm positivere Werte annahm (Abb. 3-18 C). Es bestand allerdings kein linearer Zusammenhang zwischen der NaCl-Konzentration und dem  $\beta$ -Faltblattanteil, da dieser schon ab 1 mM NaCl seinen maximalen Wert erreicht zu haben schien (Abb. 3-18 C). Aufgrund der niedrigen NaCl-Konzentrationen konnte davon davon ausgegangen, dass der erhöhte  $\beta$ -Faltblattanteil nicht durch Proteinaggregation infolge von Aussalzungsprozessen verursacht wurde, zumal cwCT(F34W) in ASW (und somit in Gegenwart von u. a. 450 mM NaCl) bei Abwesenheit von DSPG-SUV stabil in Lösung blieb und die gleiche Konformation wie in MQ aufwies (Abb. 3-13 C). Das deutete darauf hin, dass der die Induktion der  $\beta$ -Faltblattstruktur verstärkende Effekt des NaCl eher eine Folge der durch die Salzionen erfolgten Ladungsabschirmung der SUV-Oberfläche und darauf beruhender Effekte war.

Um auszuschließen, dass es sich bei der Induktion von  $\beta$ -Faltblattkonformation in cwCT(F34W) um einen vorrangig durch unspezifische hydrophobe Wechselwirkungen an der gekrümmten SUV-Lösungsmittel-Grenzfläche verursachten Effekt handelte, wurde in wässriger gepufferter Lösung befindliches cwCT(F34W) in Gegenwart von Polystyrol-Kügelchen (PSK) mit Durchmessern von  $\sim 30$  nm inkubiert. In Gegenwart der PSK glich das Fern-UV CD-Spektrum von cwCT(F34W) dem des Proteins in wässriger gepufferter Lösung. Der Verlauf beider Spektren mit ausgeprägt negativer molarer Elliptizität bei  $< 200$  nm sowie dem schwach ausgeprägten Maximum bei  $\sim 220$  nm zeigte, dass cwCT(F34W) in beiden Fällen intrinsisch entfaltet war (Abb. 3-19 A). Die UV-spektroskopische Bestimmung der cwCT(F34W)-Konzentration im Überstand nach Abzentrifugieren der PSK lieferte zudem das Ergebnis, dass die cwCT(F34W) Konzentration immer noch genauso hoch wie vor Zugabe der PSK war und das Protein folglich nicht unspezifisch an den PSK haften blieb. Da gekrümmte, hydrophobe Oberflächen, die etwa den Dimensionen der Oberfläche der aus 5 mM DSPG hergestellten SUV entsprachen, keinen Einfluss auf die Konformation von cwCT(F34W) ausübten, konnte bestätigt werden, dass die durch DSPG-SUV induzierten Strukturänderungen nicht auf unspezifischen hydrophoben Grenzflächenwechselwirkungen beruhten.



**Abb. 3-19. Fern-UV-CD-Spektren zur Untersuchung des Einflusses der Oberflächenbeschaffenheit und Lipidmembranmorphologie auf die cwCT(F34W) Sekundärstruktur.** (A) Spektren von cwCT(F34W) in wässriger gepufferter Lösung, in Gegenwart von Polystyrolkugeln (PSK) mit 30 nm Durchmesser sowie cwCT(F34W)-haltiger Überstand nach Abzentrifugieren der PSK. (B) Spektren von cwCT(F34W) in wässriger gepufferter Lösung sowie in Anwesenheit von DSPG- bzw. DSPC-SUV. (C) Spektren von cwCT(F34W) in Gegenwart von DSPG-SUV unter- und oberhalb der DSPG-Übergangstemperatur (55 °C). (D) Spektren von cwCT(F34W) in gepufferter wässriger Lösung, in Gegenwart von DSPC-SUV sowie in Gegenwart multilamellarer DSPG-Lipidstrukturen.

Da es sich bei der DSPG-SUV-vermittelten Induktion von  $\beta$ -Faltblattstruktur in cwCT(F34W) um das Resultat spezifischer Interaktionen zu handeln schien, wurde die Rolle der SUV-Oberflächenladung sowie der Lipidmembranbeschaffenheit untersucht. Zunächst wurde analysiert, welchen Einfluss die Oberflächenladung der SUV spielte. Zu diesem Zweck wurden SUV aus 1,2-distearoyl-*sn*-glycero-3-phosphocholin (DSPC) hergestellt, welches sich von DSPG nur durch seine Kopfgruppe – Cholat statt Phosphat – unterscheidet und somit kein Anion sondern ein Zwitterion darstellt. Verglichen mit den in Anwesenheit von DSPG-SUV gemessenen cwCT(F34W)-CD-Spektren wies ein in Gegenwart von DSPC-SUV aufgenommenes CD-Spektrum nicht das für eine  $\beta$ -Faltblattstruktur typische Minimum bei 218 nm auf und zeigte einen Verlauf, welcher qualitativ dem des intrinsisch entfalteten cwCT(F34W) in wässriger, gepufferter Lösung gleich (Abb. 3-19 B).

Die Anwesenheit der DSPC-SUV hatte keinen Einfluss auf die Sekundärstruktur von cwCT(F34W) und eine negative Oberflächenladung der SUV schien somit eine wichtige Rolle bei der Induktion der  $\beta$ -Faltblattkonformation zu spielen.

Wurde die Integrität der DSPG-SUV-Doppelmembran dadurch gestört, dass die SUV über die Übergangstemperatur des Lipids erwärmt wurden, war ein Einfluss auf die Sekundärstruktur von cwCT(F34W) detektierbar. Als die DSPG-SUV-cwCT(F34W)-Emulsion über die Übergangstemperatur des DSPG (55 °C) auf 70 °C erwärmt wurde, entsprach das cwCT(F34W)-Spektrum qualitativ demjenigen, welches zuvor bei 70 °C von cwCT(F34W) in wässriger, gepufferter Lösung aufgenommen wurde (Abb. 3-19 C). Dies war ein Beleg dafür, dass die  $\beta$ -Faltblattkonformation des cwCT(F34W), die durch Kontakt mit DSPG-SUV bei 20 °C induziert worden war, wieder in eine auch für lösliches cwCT(F34W) in Abwesenheit der SUV bei 70 °C vorherrschende, intrinsisch entfaltete Konformation übergegangen war. Ein zusätzliches Indiz dafür, dass die Induktion von  $\beta$ -Faltblattstruktur in cwCT(F34W) von der Morphologie der Membranstruktur abhängig zu sein schien, lieferten Versuche, bei denen cwCT(F34W) mit multilamellaren DSPG-Strukturen in Kontakt trat. Anders als in Gegenwart der DSPG-SUV war in den cwCT(F34W)-CD-Spektren kein ausgeprägtes Minimum bei ~ 218 nm sichtbar und das Spektrum glich dem von cwCT(F34W) in gepufferter, wässriger Lösung (Abb. 3-19 D). Somit hatten multilamellare DSPG-Lipidstrukturen im Gegensatz zu den DSPG-SUV mit ihrer Lipiddoppelmembranstruktur keinen Einfluss auf die Konformation von cwCT(F34W).

#### **3.7 Partitioning-folding coupling von cwCT(F34W) an Lipidmembrangrenzflächen**

Die Vermittlung von Proteinfaltung durch Lipidmembranen und damit die Entstehung von definierten Sekundärstrukturelementen innerhalb von IDPs wird als *partitioning-folding coupling* bezeichnet und führt bei Membranproteinen zur Bildung von transmembranen  $\alpha$ -Helixbündeln bzw.  $\beta$ -Fässern<sup>[206,207]</sup>. Wie die Bezeichnung dieses Vorgangs impliziert, kommt es dabei zu einer „Aufteilung“ des IDP zwischen der wässrigen Lösung und der komplexen Grenzfläche der Lipidmembran mit simultaner Faltung und Insertion des Proteins. Obwohl thermodynamische Vorhersagen über PFC aufgrund des

vielschichtigen Zusammenspiels von Aufteilung, Proteinfaltung, Membraninsertion und, in vielen Fällen, Oligomerisierung komplex sind<sup>[193]</sup>, gelang auf Grundlage experimenteller Daten inzwischen die Entwicklung eines Vorhersagealgorithmus. Dieser trifft auf Grundlage der Sequenz eines Proteins Vorhersagen bezüglich der Neigung zu PFC-Verhalten und berechnet, ob in dessen Verlauf vorzugsweise  $\alpha$ -helikale oder  $\beta$ -Faltblattstrukturen gebildet werden<sup>[207]</sup>.

Wegen der Ähnlichkeit von PFC und der in Kapitel 3.6 nachgewiesenen Induktion einer  $\beta$ -Faltblattkonformation in cwCT(F34W) nach Interaktion mit SUV wurde mit Hilfe des MPEx-Algorithmus<sup>[151]</sup> untersucht, ob es sich dabei um einen mit PFC vergleichbaren Effekt handelte. Als Kontrollen wurden zudem die Sequenzen von unverändertem cwCT (ohne Tyrosin-zu-Tryptophan Mutation), eADF4(C16) und Sfl analysiert. MPEx ermittelte für einen 61 Aminosäuren umfassenden Bereich von cwCT(F34W) (einschließlich des an Position sechs liegenden Tryptophans) eine Freie Energie von -17,71 kJ/mol bezüglich der Interaktion („Aufteilung“) von cwCT(F34W) mit der Grenzflächenschicht einer Lipiddoppelmembran (Tab. 3-2).

**Tab. 3-2. Zusammenfassung der bioinformatischen Analysen bezüglich des PFC-Verhaltens von cwCT(F34W), cwCT, eADF4(C16) und Sfl.** MPEx ermittelt Sequenzbereiche in Proteinen, die unter Berücksichtigung der Freien Enthalpie des Vorgangs aus einer wässrigen Umgebung in eine Lipidmembrangrenzfläche eindringen können und trifft Vorhersagen über die damit verbundenen Sekundärstrukturänderungen. n. b.: von MPEx nicht bestimmbar

Protein	Hydropathizitätsanalyse			$\beta$ -Struktur Analyse		
	Anzahl vorhergesagter Bereiche	Länge des AS-Abschnitts	$\Delta G$ [kJ/mol]	Anzahl der $\beta$ -Haarnadelbereiche	Länge des AS-Abschnitts	Punktzahl bzgl. der Ausbildung von $\beta$ -Faltblattkonformationen
cwCT(F34W)	1	61	-17,71	1	32	3,18
cwCT	1	71	-15,32	1	32	3,18
eADF4(C16)	16	8	-0,63	0	0	n. b.
Sfl	0	0	n. b.	3	191 12 41	7,56

Innerhalb dieses Bereichs wurde darüber hinaus eine zusammenhängende, 32 Aminosäuren umfassende Sequenz – mit Tryptophan beginnend - identifiziert, die aufgrund der Aminosäureabfolge theoretisch in der Lage ist, während des PFC eine  $\beta$ -Haarnadelstruktur einzunehmen (Tab. 3-2).

Für das nichtmutierte Gegenstück von cwCT(F34W), cwCT, lag die vorhergesagte Freie Energie (unter identischen Analyseparametern) bei -15,32 kJ/mol und damit etwas niedriger als für cwCT(F34W). Die Punktzahl und damit inhärente Fähigkeit der Aminosäuresequenz, eine  $\beta$ -Faltblattstruktur auszubilden, war jedoch identisch. Die hauptverantwortlich zeichnende Aminosäuresequenz umfasste in diesem Fall 71 Aminosäuren und innerhalb dieser Sequenz wurde wiederum ein 32 Aminosäuren umfassender Bereich (diesmal direkt mit dem Phenylalanin beginnend) identifiziert, dessen Aminosäuren ebenfalls bevorzugt eine  $\beta$ -Haarnadelkonformation einnehmen (Tab. 3-2).

Die bioinformatische Analyse für das als Kontrolle verwendete Spinnenseidenprotein eADF4(C16) führte zur Identifizierung von 16 identischen, je acht Aminosäuren umfassenden Sequenzbereichen, deren Hydropathizität theoretisch für Interaktionen des Proteins mit der Grenzfläche von Lipidmembranen in Frage kam. Die Länge jedes einzelnen Bereichs war jedoch deutlich kürzer, als dies bei cwCT(F34W) oder cwCT der Fall war. Die errechnete Freie Energie lag mit nur -0,63 kJ/mol um etwa den Faktor 30 unter der für cwCT(F34W) ermittelten. Darüber hinaus wurden auch keine Sequenzen erkannt, die im Hinblick auf PFC ein inhärentes Potential zur Ausbildung von  $\beta$ -Faltblattstrukturen besitzen würden (Tab. 3-2). Im zweiten als Kontrolle verwendeten Spinnenseidenprotein Sfl konnten durch MPEx drei Sequenzbereiche identifiziert werden, die theoretisch in der Lage wären,  $\beta$ -Faltblattstrukturen auszubilden. Die dafür errechnete Punktzahl war sogar nahezu doppelt so hoch wie für die entsprechenden cwCT(F34W)-Sequenzen. Es konnte aber kein Sequenzbereich ermittelt werden, dessen Hydropathizität eine Interaktion von Sfl mit der Grenzfläche von Lipiddoppelmembranen wahrscheinlich macht (Tab. 3-2). Das Ergebnis der hier durchgeführten bioinformatischen Analyse war, dass es sich bei der DSPG-SUV induzierten Bildung von  $\beta$ -Faltblattstrukturen bei cwCT(F34W) höchstwahrscheinlich um einen PFC-ähnlichen Effekt handelte, während dies bei den weiteren untersuchten Analoga extrakorporaler Strukturproteine (eADF4(C16) und Sfl) nicht der Fall gewesen zu sein schien (vgl. Kapitel 3.6.2).

## 4 Diskussion

### 4.1 Rekombinante Herstellung der flankierenden Domänen der *M. galloprovincialis* preCols

#### 4.1.1 Unterschiede zwischen den zur rekombinanten Herstellung der *flank*-Proteine verwendeten cDNA-Sequenzen und bereits bekannten Isoformen

Für aus der cDNA-Bank (Mfl-1) des Lehrstuhls Biomaterialien von Dr. Michael Suhre und Adrian Golser isolierte *M. galloprovincialis* preCol-Sequenzen wurde bereits gezeigt<sup>[125]</sup>, dass diese bezogen auf bereits veröffentlichte Sequenzen<sup>[109]</sup> Isoformen der entsprechenden Proteine codieren.

Sequenzuntersuchungen der zur Verfügung gestellten preCol-Klone, die im Rahmen dieser Arbeit zur Erzeugung der für die rekombinante Produktion von *M. galloprovincialis* *flank*-Proteinen benötigten Vektorkonstrukte benutzt wurden, zeigten hinsichtlich der flankierenden Regionen sowohl für den verwendeten preCol-D als auch den verwendeten preCol-NG Klon weitere Sequenzunterschiede im Vergleich zu den am Lehrstuhl Biomaterialien schon bekannten Isoformen<sup>[125]</sup>. Die Aminosäuresequenz der carboxyterminal flankierenden Region des preCol-D glich zu 97,2 % und die der aminoterminal flankierenden Region des preCol-NG zu 98,2 % den entsprechenden, bereits aus der Mfl-1 bekannten Sequenzen. Die Unterschiede bestanden im Fall von preCol-D aus einer GGA- anstelle einer GAG-Insertion<sup>[125]</sup> (Abb. 3-1), was dazu führt, dass sich in der seidenähnlichen flankierenden Region anstelle eines (AG)<sub>2</sub>-Konsensusmotivs ein ebenfalls für diesen Sequenzbereich charakteristisches und ebenfalls häufig auftretendes AGG-Konsensusmotiv wiederfindet (Tab. 1-1)<sup>[31,57]</sup>. Die im Fall der aminoterminal flankierenden Region des preCol-NG anstelle der bereits bekannten GGGL-Insertion<sup>[125]</sup> vorliegende GLGG-Insertion (Abb. 3-1) führt lediglich zu einer Veränderung der Zahl an Glycinen innerhalb eines für diese Domäne charakteristischen (XG<sub>m</sub>)<sub>n</sub>-Konsensusmotivs (Tab. 1-1 & Abb. 1-5)<sup>[31,58]</sup>. Somit besaßen auch die im Rahmen dieser Arbeit verwendeten

Isoformen der flankierenden Regionen von preCol-D und preCol-NG den für sie jeweils typischen seiden- bzw. pflanzenzellwandähnlichen Charakter.

### 4.1.2 Rekombinante Herstellung von *flank*-Proteinen in *E. coli*

Als Wirtsorganismus für die rekombinante Herstellung der *M. galloprovincialis flank*-Proteine wurde *E. coli* verwendet. Der Vorteil der Verwendung von *E. coli*-Zellen liegt darin, dass deren genetischer Hintergrund bereits umfassend erforscht ist<sup>[208,209]</sup> und dass diese in standardmäßig eingerichteten molekularbiologischen Laboratorien ohne größeren Aufwand gehandhabt werden können<sup>[210,211]</sup>. Die Verfügbarkeit eines breiten Spektrums genetischer und molekularbiologischer Methoden zur Manipulierung von *E. coli*<sup>[211,212]</sup> ermöglicht das schnelle und effiziente Einbringen transgener Sequenzen, was in Kombination mit den geringen Generationszeiten<sup>[210,211]</sup> den zeitlichen Aufwand für das Auffinden geeigneter Bedingungen, die zur rekombinanten Herstellung der Zielproteine benötigt werden, verkürzt. So können innerhalb eines kürzeren Zeitraums (verglichen mit anderen eukaryotischen Expressionssystemen) eine Vielzahl rekombinanter Proteine (bzw. Varianten eines Proteins) hergestellt, analysiert und untereinander verglichen werden.

Die rekombinante Produktion von Proteinen eukaryotischen Ursprungs in *E. coli* ist jedoch nicht immer unproblematisch, da den Bakterien eukaryotische Chaperone, Systeme zur posttranslationalen Modifikation von Proteinen und eine Zellkompartimentierung fehlen, was zu Fehlfaltungen und zur Aggregation des heterologen Proteins während dessen Biosynthese führen kann<sup>[210,211]</sup>. Da es sich bei den rekombinant herzustellenden *flank*-Proteinen jedoch um niedermolekulare Proteine handelt, deren korrespondierende Aminosäurebereiche innerhalb der preCols keine posttranslationalen Modifikationen aufweisen, und auch schon zahlreiche andere eukaryotische Proteine<sup>[211,213]</sup>, darunter zahlreiche Spinnen- und Insektenseidenproteine (die wie die preCols ebenfalls Bestandteil schwer löslicher, extrakorporaler Proteinstrukturen sind) sowie davon abgeleitete artifizielle Proteine<sup>[148,214-219]</sup>, erfolgreich in *E. coli* hergestellt werden konnten, wurde dieser Wirtsorganismus zur Herstellung der *flank*-Proteine verwendet.

Bei den unveränderten *flank*-Proteinen mit aminoterminalen (His)<sub>6</sub>-tag (vgl. Abb. 3-4 & 3.2.1) konnte nur für (His)<sub>6</sub>-sCT, -eCT und -cwCT eine rekombinante Produktion in *E. coli*



nachgewiesen werden, jedoch war auch dieses Resultat nicht verlässlich reproduzierbar (vgl. Kapitel 3.2.1). Die synthetisierten Proteine befanden sich nach der Zelllyse fast ausschließlich in der unlöslichen Fraktion. Auf Grund der Tatsache, dass *in silico*-Sekundärstrukturvorhersagen für o. g. *flank*-Proteine (Abb. 3-2) auf eine zumindest partielle Faltung hinwiesen, war das Ziel jedoch, rekombinant produzierte *flank*-Proteine aus der löslichen Zellfraktion zu reinigen. Auf diese Weise können später nötig werdende Um- bzw. Rückfaltungsprozesse, die mit der Gefahr von Fehlfaltungen verbunden sein könnten, vermieden und durch die Charakterisierung der *flank*-Proteine ein möglichst exaktes Bild über die native Konformation sowie Funktion/Rolle der flankierenden Domänen der preCols erhalten werden.

Der zur Verbesserung der Löslichkeitseigenschaften verwendete aminoterminal SUMO-*tag*<sup>[169,196,197]</sup> zeigte den gewünschten Einfluss, auch wenn sich im Fall der SUMO-*flank*-Fusionsproteine weiterhin ein kleinerer Teil der Fusionsproteine in der unlöslichen Fraktion befand (Abb. 3-6). Neben den jeweiligen Volllängenproteinen waren sowohl in während als auch in nach der Expression genommenen Proben verkürzte Proteinfragmente derselben nachweisbar (Abb. 3-6). Als mögliche Gründe kommen eine schlechte Translationseffizienz der hierfür verwendeten eukaryotischen *M. galloprovincialis* Nukleotidsequenzen und daraus resultierende vorzeitige Translationsabbrüche sowie *in vivo* auftretende proteolytische Abbauprozesse in Frage. Speziesbedingte Unterschiede in der Codonverwendung (d. h. die Häufigkeit mit der synonyme DNA-Codons zur Codierung einer Aminosäure verwendet werden) nehmen Einfluss auf Genauigkeit und Effizienz der Translation und beeinflussen die heterologe Herstellung rekombinanter Proteine dadurch entscheidend<sup>[157,220-224]</sup>. Da im Rahmen dieser Arbeit zur Herstellung von *flank*-Proteinen jedoch auch an die *E. coli* Codonverwendung angepasste Nukleotidsequenzen verwendet wurden, nach deren Expression und Translation ebenfalls verkürzte Fragmente der entsprechenden rekombinanten *flank*-Fusionsproteine nachweisbar waren (Abb. 3-7), schienen Unterschiede in der Codonverwendung in diesem Fall eine untergeordnete Rolle zu spielen. Somit schien es weitaus wahrscheinlicher, dass die Entstehung der Fragmente das Ergebnis eines verstärkten und bereits während der rekombinanten Produktion der *flank*-Fusionsproteine auftretenden proteolytischen Abbauprozesses war, der mit der Primärstruktur der *flank*-Proteine in Zusammenhang stand. Deren Primärstruktur erfüllt wegen ihres hochrepetitiven und auf wenige verwendete

Aminosäuren beschränkten Aufbaus Kriterien von sog. „Regionen niedriger Komplexität“<sup>[225]</sup>. Dabei handelt es sich um Sequenzabschnitte innerhalb von Proteinen, die reich an bestimmten Aminosäuren sind (relevant für die *flank*-Sequenzen sind v. a. A, G, S, L und F, Abb. 3-3), welche sowohl unregelmäßig wie auch regelmäßig als kurze Wiederholungseinheiten angeordnet sein können<sup>[225]</sup>. Für rekombinant in *E. coli* produzierte Proteine, die solche Regionen niedriger Komplexität umfassen, konnte bereits eine schlechtere Löslichkeit nachgewiesen werden. Zudem werden Regionen niedriger Komplexität auch für verstärkt auftretende *in vivo* Degradation verantwortlich gemacht<sup>[211]</sup>.

Expressionstests eines preCol-NG *flank*-Konstrukts, welches neben dem aminoterminalen SUMO-*tag* zusätzlich für eine carboxyterminale Chitin-Bindedomäne (CBD-*tag*) codierte (SUMO-cwCT(F34W)-CBD), zeigten, dass die bei den SUMO-*flank*-Fusionsproteinen beobachtbaren proteolytischen Abbauprozesse bei einer Abschirmung der eigentlichen *flank*-Aminosäuresequenz mittels eines zweiten, globulären *tags* nicht mehr nachweisbar waren und eine *in vivo* Degradation durch die Abschirmung unterbunden wurde. Während der chromatographischen Reinigung von cwCT(F34W) konnte neben SUMO-cwCT(F34W)-CBD zwar auch noch SUMO-cwCT(F34W) nachgewiesen werden (Abb. 3-10 C & D), jedoch wurde dieses nicht infolge enzymatische Abbauprozesse gebildet, sondern entstand einerseits durch die Abspaltung des CBD-*tags* während der SDS-PAGE Analyse der Proben, die vom im SDS-Auftragspuffer enthaltenen 2-Mercaptoethanol induziert wurde, sowie andererseits auch durch schon *in vivo* induzierte Abspaltung des CBD-*tags*<sup>[226-228]</sup>. Eine weitere Verbesserung der Löslichkeitseigenschaften durch die Verwendung des CBD-*tags* konnte – bezogen auf die SUMO-*flank*-Fusionsproteine – nicht beobachtet werden.

### 4.1.3 Reinigung rekombinant hergestellter *flank*-Proteine

Unter den Bedingungen, bei welchen während der rekombinanten Produktion von *flank*-Proteinen neben dem Volllängenprotein auch verkürzte Fragmente gebildet wurden (vgl. Kapitel 3.2.2), war eine zufriedenstellende Reinigung des jeweiligen Volllängenproteins nicht möglich. Die verschiedenen *flank*-Proteine konnten über den (His)<sub>6</sub>-*tag* zwar per IMAC von den *E. coli*-eigenen Proteinen, aber nicht von den verkürzten Fragmenten, abgetrennt werden, da diese ebenfalls einen (His)<sub>6</sub>-*tag* trugen. Da die Fragmente

darüber hinaus sehr ähnliche physikalisch-chemische Eigenschaften wie das Volllängenprotein besitzen und die Molekulargewichtsunterschiede zum Volllängenprotein oft nur gering ausfielen, gelang eine Optimierung der Reinheit des Volllängenproteins auch durch Gelpermeationschromatographie nicht.

Erst die Herstellung der *flank*-Proteine in Form von Fusionsproteinen mit zwei strukturierten, terminalen *tags*, bei denen *in vivo* keine Degradation zu beobachten war (Abb. 3-7 & Kapitel 3.2.2), ermöglichte, wie am Beispiel von SUMO-cwCT(F34W)-CBD gezeigt, die Entwicklung einer Methode zur Reinigung des rekombinant erzeugten *M. galloprovincialis flank*-Proteins (vgl. Kapitel 3.4). Durch die am aminoterminalen SUMO-*tag* vorhandene (His)<sub>6</sub>-Sequenz konnte SUMO-cwCT(F34W)-CBD per IMAC von den *E. coli*-Proteinen abgetrennt werden. Während die Ulp1-vermittelte Abspaltung des SUMO-*tags* quantitativ möglich war, war eine quantitative Abspaltung des CBD-*tags* durch DTT-induziertes Spleißen nicht möglich. *In vitro* wird die Effizienz des autokatalytischen Spleißvorgangs neben der Temperatur und der Inkubationsdauer entscheidend auch durch die Aminosäure, welche sich direkt *upstream* der Inteinsequenz befindet, beeinflusst. Für das in der cwCT(F34W)-Sequenz an dieser Position auftretende Ala beträgt die Spleißeffizienz bei 4 °C und 40 h Inkubationsdauer, was den im Rahmen dieser Arbeit verwendeten Bedingungen von 6 °C und 48 h recht nahe kommt, 75-90 %<sup>[226,228]</sup>. Dieser Wert stellt gleichzeitig das Optimum der Spleißeffizienz des verwendeten IMPACT™-Systems bei 6 °C dar<sup>[228]</sup>. Eine leichte Verbesserung kann noch über Erhöhung der Temperatur erreicht werden<sup>[228]</sup>, auch wenn eine vollständig quantitative Abspaltung des CBD-*tags* nicht möglich scheint. Das durch unvollständiges Spleißen noch verbliebene cwCT(F34W)-CBD konnte während der abschließenden GPC vollständig von cwCT(F34W) abgetrennt werden (Abb. 3-11), die nicht quantitative Spaltung führte jedoch insgesamt zu einer Verminderung der cwCT(F34W)-Ausbeute während der Reinigung.

Die exemplarisch an SUMO-cwCT(F34W)-CBD erarbeitete Strategie zur rekombinanten Produktion und Reinigung eines *M. galloprovincialis flank*-Proteins in *E. coli* sollte auch die Herstellung und Reinigung der anderen preCol-*flanks* über die entsprechenden SUMO-CBD-Fusionsproteine ermöglichen. Da die Effizienz des zur CBD-*tag* Abspaltung induzierten Spleißvorgangs, die einen signifikanten Einfluss auf die Ausbeute während der Reinigung besitzt, jedoch stark von der carboxyterminalen Aminosäure des

*flank*-Proteins beeinflusst wird, wäre im Fall der aminoterminal flankierenden Regionen aller drei preCols (eNT, sNT und cwNT, Abb. 3-4 & Kapitel 11.1.1) eine Erweiterung der auf Basis von Lucas *et al.* für diese Arbeit festgelegten Domänengrenzen nötig, um für die CBD-*tag* Abspaltung günstige Aminosäuren an dieser Position zu erhalten. Im Fall von cwNT genügt die Erweiterung der Sequenz um eine Aminosäure, während die Sequenzen von eNT und sNT um ein GPG bzw. GGP-Motiv erweitert und die carboxyterminale Aminosäure zusätzlich zu Alanin mutiert werden müsste. In beiden Fällen würde das carboxyterminale Alanin für die bestmögliche Spleißeffizienz sorgen, ohne jedoch den für die jeweilige flankierende Region typische Motivcharakter zu verändern, da es sich ausschließlich um die Vervollständigung eines elastinähnlichen  $X^1X^2X^3PX^4$  (eNT) bzw. seidenähnlichen XGG-Motivs (sNT) handeln würde (vgl. Kapitel 1.3.2.2).

### 4.2 Struktur und Stabilität des preCol-NG *flank*-Proteins cwCT(F34W) in Lösung

Gereinigtes, lyophilisiertes cwCT(F34W) konnte in verschiedenen Lösungsmitteln komplett gelöst und so seine Struktur wie auch Stabilität in Lösung mittels spektroskopischer Methoden charakterisiert werden.

UV-Spektren von in MQ, AMW sowie gepufferter wässriger Lösung gelöstem cwCT(F34W) wiesen eine stabile Nulllinie oberhalb einer Wellenlänge von 320 nm auf (Abb. 3-13 A), was bedeutete, dass keine durch eventuell in der Lösung befindlichen Aggregate verursachte Streuung auftrat und sich das Protein komplett gelöst hatte. Nach einstündiger Inkubation bei Raumtemperatur und anschließender Ultrazentrifugation glichen zudem sowohl die UV-Spektren wie auch die Absorptionswerte bei 280 nm denen von frisch gelöstem cwCT(F34W) (Abb. 3-13 A), was zeigte, dass das Protein unter den genannten Bedingungen innerhalb dieses Zeitraums komplett in Lösung blieb.

Qualitative Aussagen über die Sekundärstruktur des gelösten cwCT(F34W) konnten anhand von Fern-UV CD-Spektren (< 250 nm) gemacht werden, da die in diesem Spektralbereich absorbierenden Peptidbindungen abhängig von der Konformation des Proteinerückgrats die durch den chiralen Charakter der L-Aminosäuren hervorgerufenen CD-Absorptionsbanden zusätzlich beeinflussen<sup>[185,229]</sup>. Fern-UV CD-Spektren von in MQ als auch von in ASW gelöstem cwCT(F34W) (Abb. 3-13 C) zeigten, dass das Protein eine

intrinsisch entfaltete Konformation besaß<sup>[184,185]</sup>. Dieses Ergebnis wurde auch von den Strukturuntersuchungen zum thermischen Verhalten von cwCT(F34W) gestützt, da Proteine mit hohem Sekundärstrukturanteil häufig thermisch entfalten, was oft auch mit einer Verstärkung intermolekularer hydrophober Interaktionen und der Aggregation der Proteine einhergeht. Solch ein Verhalten war für cwCT(F34W) während des Aufheizens jedoch nicht nachweisbar und die detektierten CD-Spektren ließen weder auf strukturelle Änderungen noch Proteinaggregation schließen (Abb. 3-14).

Ein schwach ausgeprägtes, breites CD-Maximum bei ~ 215 nm deutete zudem darauf hin, dass neben der hauptsächlich intrinsisch entfalteten Konformation in untergeordnetem Maße auch  $3_1$ -helikale Strukturanteile vorhanden waren. Diese treten häufig in glycinreichen Proteinen auf und konnten auch schon in verschiedenen anderen Proteinen mit ähnlich glycinreicher Aminosäurezusammensetzung nachgewiesen werden<sup>[198,230]</sup>. Solch eine im Fall glycinreicher Proteine auch als Polyglycin-II (PGII-) Konformation bezeichnete Struktur ist durch linksgängige Helices charakterisiert, die den gestreckten und ebenfalls linksgängigen  $3_1$ -Helices der Polyprolin-II (PPII-) Konformation ähneln<sup>[230,231]</sup>. Die gleichen sterischen Gründe wie bei letzterer führen auch im Fall von PGII-Helices dazu, dass sowohl die Carbonylgruppen des Proteinrückgrats als auch die Aminosäureseitenketten nach außen zeigen und deshalb gut mit Lösungsmittel- sowie anderen Molekülen über Wasserstoffbrückenbindungen interagieren können<sup>[231]</sup>. Die vorteilhafte Interaktion mit Lösungsmittelmolekülen ist auch der Hauptgrund, warum es gerade in IDP häufig zur Ausbildung dieser  $3_1$ -helikalen Strukturanteile kommt<sup>[230]</sup>.

Die durch Fern-UV CD-Spektroskopie nachgewiesene Sekundärstruktur von cwCT(F34W) spiegelte sich auch im Fluoreszenzemissionsspektrum des an Position 34 befindlichen Tryptophans wieder (Abb. 3-13 B), da die Breite des Emissionspeaks in Kombination mit dem Maximum bei ~ 354 nm typisch für ein Tryptophanemissionsspektrum der Spektralform III, welches im Allgemeinen durch sich an der Proteinoberfläche in Kontakt mit freien Wassermolekülen befindliche Tryptophanreste hervorgerufen wird<sup>[189]</sup>, ist. Die für rekombinantes cwCT(F34W) qualitativ ermittelte Sekundärstruktur deckt sich, auch wenn die anderen preCol-Domänen fehlen, mit der in der Literatur beschriebenen amorphen und weitgehend unstrukturierten Morphologie der flankierenden Regionen von *M. californianus* preCol-NG<sup>[94]</sup>. Da in Lösung befindliches cwCT(F34W) auch ohne die His/DOPA-Termini und die zentrale Kollagendomäne des preCol-NG eine den natürlich

vorkommenden flankierenden Regionen vergleichbare Struktur einnimmt, kann davon ausgegangen werden, dass letztere auch *in vivo* (zumindest solange preCol-NG in gelöstem Zustand vorliegt) ihre Konformation relativ unabhängig von den anderen Domänen einnehmen.

Weiterhin ist es wahrscheinlich, dass das intrinsisch entfaltete cwCT(F34W) aufgrund seiner Aminosäuresequenz, die durch die Abwesenheit stark hydrophober sowie einer größeren Anzahl geladener Aminosäuren gekennzeichnet ist, zu einer durch das Proteinerückgrat vermittelten Kompaktierung neigt<sup>[232,233]</sup>. Solch eine Kompaktierung einer intrinsisch entfalteten Struktur wurde bereits für das kleinere der beiden Anti-Frost-Proteine des Schneeflohs *Hypogastrura harveyi* Folsom<sup>[234]</sup>, welches eine dem cwCT(F34W) ähnliche Primärstruktur besitzt, nachgewiesen<sup>[235]</sup> und könnte ein Grund dafür sein, weshalb die terminalen Regionen des *M. galloprovincialis* preCol-NG in rasterkraftmikroskopischen Aufnahmen nativer Byssusfäden eine kompakte, globuläre Morphologie aufweisen<sup>[59]</sup>.

Das für die flankierenden Regionen von preCol-NG angenommene<sup>[31,58]</sup> und teilweise experimentell nachgewiesene<sup>[73]</sup> Auftreten von  $\beta$ -Faltblattstrukturen konnte im Fall von cwCT(F34W) nicht bestätigt werden, obwohl dies wegen der Ähnlichkeit seiner (XG<sub>m</sub>)<sub>n</sub>-Konsensusmotive zu den (XG)<sub>n</sub>-Motiven der Arthropodenseidenproteine sowie v. a. zu den XG<sub>n</sub>-Motiven der in pflanzlichen Zellwänden vorkommenden glycinreichen Proteine, welche beide  $\beta$ -Faltblattstrukturen ausbilden können<sup>[3,20,70-72,74,75,130,131]</sup> (Tab. 1-1), naheliegend wäre. Die Erklärung für diese scheinbar unterschiedlichen Ergebnisse findet sich jedoch in der experimentellen Auslegung, mithilfe derer die o. g.  $\beta$ -Faltblattkonformation der flankierenden Regionen von preCol-NG ermittelt wurde<sup>[73]</sup>. Da diese weder an isolierten preCol-NG Molekülen, geschweige denn an isolierten flankierenden Regionen, sondern an nativen Byssusfäden nachgewiesen wurde, macht es der Umstand, dass Byssusfäden aus mehreren verschiedenen Proteinen aufgebaut sind<sup>[41,43]</sup>, nahezu unmöglich, bestimmte Sekundärstrukturelemente mit absoluter Sicherheit einem der beteiligten Proteine oder gar einer spezifischen Domäne innerhalb dieser Proteine zuzuweisen.

### 4.3 Interaktion von cwCT(F34W) mit Mizellen und kleinen unilamellaren Vesikeln (SUVs)

Intrinsisch entfaltete Proteine (IDPs), zu denen auch die lösliche Form des *flank*-Proteins cwCT(F34W) gezählt werden kann, können wegen ihrer konformationellen Flexibilität oft mehrere Funktionen erfüllen, indem abhängig vom Interaktionspartner (i. d. R. andere Moleküle oder makromolekulare Strukturen) unterschiedliche Strukturen eingenommen werden können, wobei das IDP erst dann seine für die jeweilige Funktion nötige aktive Konformation einnimmt<sup>[199-201]</sup>. Ein Typ makromolekularer Strukturen, die als Interaktionspartner von IDPs fungieren können, sind Lipidvesikel bzw. -membranen. IDPs, die mit Lipidstrukturen interagieren, können dabei abhängig von ihren intrinsischen Eigenschaften sowohl in helikale wie auch  $\beta$ -Faltblattkonformationen falten<sup>[236-240]</sup>.

Die strukturellen Untersuchungen an löslichem cwCT(F34W), welches der carboxyterminal flankierenden Domäne des *M. galloprovincialis* preCol-NG entspricht<sup>[31,58,241]</sup>, lieferten Hinweise darauf, dass letztere — zumindest in der präassemblierten Form des preCol-Moleküls — eine intrinsisch entfaltete Region (IDR, von engl. *intrinsically disordered region*) innerhalb desselben darstellt. Da diese IDR sowohl während der Sekretion der preCols in membranumschlossenen sekretorischen Granulae als auch später im assemblierten Byssusfaden mit Lipiden und auch Wachsen in Kontakt kommen kann<sup>[37,43,123,204,205]</sup>, wurde analysiert, ob Lipidstrukturen als Interaktionspartner von cwCT(F34W) fungieren können und wie sich das *flank*-Protein nach Kontakt mit diesen verhält.

#### 4.3.1 Interaktion von cwCT(F34W) mit SDS-Mizellen

Neben seiner klassischen Rolle als Proteindenaturierungsmittel kommt SDS in Mizellenform auch als stark vereinfachtes Modellsystem bei der Nachahmung von Lipidmizellen zum Einsatz<sup>[191]</sup>. Deshalb wurden SDS-Mizellen in einem ersten Ansatz dazu verwendet, mögliche Interaktionen des intrinsisch entfalteten cwCT(F34W) mit lipidähnlichen Strukturen sowie daraus resultierende Strukturänderungen des *flank*-Proteins CD-spektroskopisch zu untersuchen. Im untersuchten SDS-Konzentrationsbereich von 1 mM

bis 50 mM (Monomerkonzentration) konnte jedoch kein strukturinduzierender Effekt des SDS auf cwCT(F34W) nachgewiesen werden (vgl. Kapitel 3.6.1). Die als Kontrollproteine verwendeten Spinnenseidenproteinanaloga eADF4(C16) und Sfl, die wie cwCT(F34W) in Lösung ebenfalls intrinsisch entfaltet sind<sup>[148,149]</sup>, zeigen in Gegenwart von SDS-Mizellen hingegen einen sehr schnellen Sekundärstrukturübergang (vgl. Kapitel 3.6.1). Die durch zwei Minima bei 208 nm sowie 222 nm geprägten CD-Spektren belegten, dass bei beiden Spinnenseidenproteinanaloga ein Strukturübergang zu einer  $\alpha$ -helikal geprägten Konformation stattfand, denn das Vorhandensein eines CD-Minimums bei 222 nm ist selbst dann ein qualitativ zuverlässiger Indikator für die Anwesenheit von  $\alpha$ -helikaler Sekundärstruktur, wenn es durch partikelbedingte Lichtstreuung (wie in Gegenwart der SDS-Mizellen) zu einer Verschlechterung des Signal-Rausch-Verhältnisses kommt<sup>[242]</sup>. Eine wie im Fall der Spinnenseidenproteinanaloga durch SDS-Mizellen vermittelte Induktion  $\alpha$ -helikaler Konformationsanteile konnte auch schon bei anderen natürlicherweise als IDP vorkommenden Proteinen wie  $\alpha$ -Synuclein,  $\beta$ -Synuclein oder dem Hefe-Hitzeschockprotein Hsp12 beobachtet werden<sup>[191,238]</sup>. Interessanterweise zeigte cwCT(F34W) also im Gegensatz zu anderen IDP und vor allem auch im Gegensatz zu den ebenfalls aus extrakorporalen Strukturproteinen abgeleiteten sowie durch ihre hochrepetitiven, glycinreichen Aminosäuresequenzen dem cwCT(F34W) sehr ähnlichen Spinnenseidenproteinen eADF4(C16) und Sfl keine Sekundärstrukturänderungen nach Kontakt mit SDS-Mizellen.

Trotz dieser Unterschiede lässt sich eine prinzipielle Interaktion von cwCT(F34W) mit Lipidstrukturen aber nicht ausschließen, da die verwendeten SDS-Mizellen ein stark vereinfachtes Lipidmodellsystem darstellen, das die biochemische Natur von Lipiddoppelmembranen nur schlecht widerspiegeln kann.

### **4.3.2 Interaktion von cwCT(F34W) mit kleinen unilamellaren Lipidvesikeln (SUV)**

#### *4.3.2.1 Induktion von Sekundärstruktur in cwCT(F34W)*

In der Natur vorkommende Lipidmembranen besitzen deutlich komplexere physikalisch-chemische Eigenschaften als die für die initialen Untersuchungen verwendeten SDS-Mizellen. Die zugrunde liegenden Phospholipidmoleküle besitzen komplexere polare

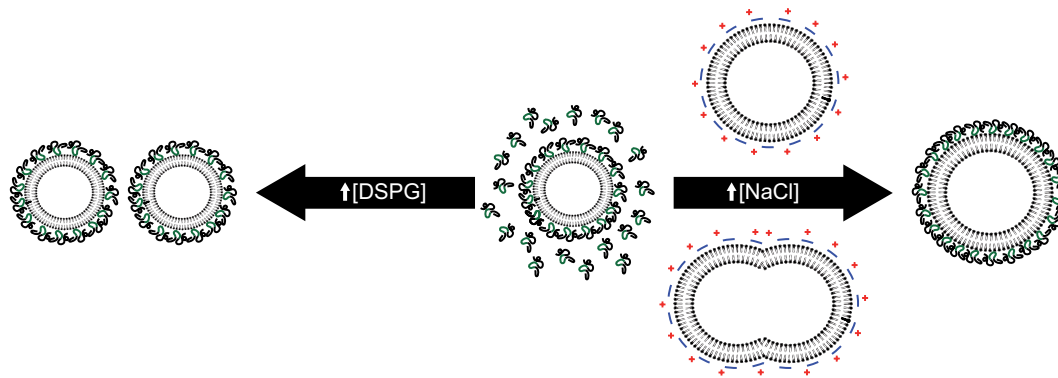


Kopfgruppen sowie unterschiedliche aliphatische Seitenketten, was in wässriger Umgebung zur Bildung einer Doppelmembran bestehend aus einem  $\sim 3$  nm dicken hydrophoben Kohlenwasserstoffkern und zweier, diesen zur wässrigen Umgebung abgrenzenden,  $\sim 1,5$  nm dicken Grenzflächenschichten führt. Innerhalb der beiden Grenzflächenschichten existiert ein steiler Gradient der Umgebungspolarität von den polaren Kopfgruppen hin zum völlig unpolaren Kohlenwasserstoffkern<sup>[193,243]</sup>. Da diese Bedingungen durch SUV exakter nachgeahmt werden können als durch SDS-Mizellen, eignen sich erstere besser zur Aufklärung der Frage, ob und wie cwCT(F34W) und damit auch die carboxy-terminal flankierende Region des preCol-NGs mit Lipidstrukturen, wie sie *in vivo* sowohl während der Sekretion der preCols als auch im assemblierten Byssusfaden vorkommen<sup>[37,43,123,204,205]</sup>, interagiert.

Die in Lösung intrinsisch entfaltet vorliegenden Proteine cwCT(F34W) sowie eADF4(C16) zeigten im Gegensatz zum ebenfalls intrinsisch entfalteten Spinnenseidenproteinanalogon Sfl sowie dem nativ gefalteten BSA in Gegenwart bestimmter SUV Änderungen ihrer Sekundärstruktur hin zu einer  $\beta$ -Faltblattkonformation (Abb. 3-16). Diese Sekundärstrukturänderungen waren das Ergebnis spezifischer Interaktionen, die sowohl von den intrinsischen Eigenschaften des Proteins als auch der SUV abhingen. Dies äußerte sich v.a. darin, dass die Induktion einer  $\beta$ -Faltblattkonformation in cwCT(F34W) und eADF4(C16) (I) die Gegenwart von SUV voraussetzte aber nicht durch gekrümmte hydrophobe Oberflächen oder multilamellare Lipidstrukturen hervorgerufen wurde (Abb. 3-19 A & D), (II) von der Lipidzusammensetzung der SUV und somit sowohl vom physikalisch-chemischen Umfeld der Grenzflächenschichten wie auch des Kohlenwasserstoffkerns abhing (Abb. 3-16 & Abb. 3-19 B) und (III) SUV des gleichen Phospholipids nicht in allen verwendeten Proteinen Sekundärstrukturänderungen induzieren konnten (vgl. Abb. 3-16 C & D).

Der Anteil an  $\beta$ -Faltblattkonformation, der durch DSPG-SUV in cwCT(F34W) induziert wurde, schien mit der zur Herstellung der SUV verwendeten DSPG-Monomerkonzentration sowie mit der Ionenstärke in direktem Zusammenhang zu stehen (Abb. 3-18). Berücksichtigt man, dass cwCT(F34W) mit einer Konzentration von  $15 \mu\text{M}$  in allen verwendeten Versuchsansätzen in einem  $\sim 100$ -fach molaren Überschuss (bezogen auf die SUV, vgl. Kapitel 2.10) vorhanden war, ist davon auszugehen, dass die CD-spektroskopisch detektierbare Erhöhung des  $\beta$ -Faltblattanteils dadurch zustande

kommt, dass höhere DSPG-Monomer Konzentrationen bzw. die Erhöhung der Ionenstärke dazu führen, dass eine größere Zahl von cwCT(F34W)-Molekülen eine  $\beta$ -Faltblattkonformation einnehmen kann, sich der  $\beta$ -Faltblattanteil pro cwCT(F34W)-Molekül jedoch nicht vergrößert.



**Abb. 4-1: Einfluss von DSPG-Monomer- und NaCl-Konzentration auf die SUV-Morphologie und die Induktion von  $\beta$ -Faltblattstrukturen in cwCT(F34W).** In Bereichen oberhalb der CMC führt eine Erhöhung der DSPG-Monomerkonzentration auch zu einer Erhöhung der Anzahl der SUV (linker Pfeil). Eine Erhöhung der Ionenstärke (rechter Pfeil), z. B. durch Zugabe von NaCl, bedingt bei doppelkettigen Lipiden eine Absenkung der CMC sowie eine Erhöhung der Aggregationszahl. Bei gleichbleibender DSPG-Monomerkonzentration hat dies mehr und auch größere SUV zur Folge. Des Weiteren führt die Anwesenheit der  $\text{Na}^+$ -Kationen zu einer Abschirmung der negativen Kopfgruppenladungen der DSPG-SUV, was eine Fusion der SUV begünstigt und so ebenfalls zur Entstehung größerer SUV beiträgt. Gemeinsam ist beiden Fällen, dass den cwCT(F34W)-Molekülen so eine vergrößerte Interaktionsfläche zur Verfügung steht.

Sowohl die Erhöhung der DSPG-Monomerkonzentration als auch eine Erhöhung der Ionenstärke besitzen einen direkten Einfluss auf die SUV-Morphologie und dadurch auf die für die cwCT(F34W)-Moleküle im ersten Schritt zugängliche Oberfläche der äußeren Grenzflächenschicht. Da sich unter gleichbleibenden Matrix- und Umgebungsbedingungen die CMC eines Amphiphils nicht ändert, führt die Erhöhung der Konzentration des Monomers in erster Linie zur Bildung einer größeren Anzahl von Vesikeln (vgl. Abb. 4-1, linke Seite). Darüber hinaus ist bekannt, dass eine Erhöhung der Ionenstärke bei Lipiden mit zwei aliphatischen Seitenketten zu einer Verringerung der CMC sowie zu einer Erhöhung der Aggregationsnummer beiträgt<sup>[244-246]</sup>, so dass auch im Fall der DSPG-SUV die Zugabe von NaCl die Bildung einer größeren Zahl sowie insgesamt größere Vesikel zur Folge hatte (Abb. 4-1, rechte Seite). Sowohl mehr als auch größere SUV bedingen eine Vergrößerung der zugänglichen Oberfläche, was mehr cwCT(F34W)-Molekülen eine

Interaktion mit der komplexen SUV-Grenzflächenregion erlaubte und sich durch scheinbar höhere  $\beta$ -Faltblattanteile in den über die Gesamtheit der Proteinmoleküle gemittelten CD-Spektren äußerte. Die Tatsache, dass in KMW und damit auch in Gegenwart von 450 mM NaCl befindliches cwCT(F34W) in Abwesenheit von SUV keine  $\beta$ -Faltblattstruktur zeigte (Abb. 3-13 C), verdeutlicht, dass die in Gegenwart von NaCl und DSPG-SUV beobachtete scheinbare Verstärkung der  $\beta$ -faltblattstrukturinduzierenden Wirkung mit einem Einfluss des NaCl auf die SUV, und nicht mit einem Einfluss auf cwCT(F34W) selbst, in Zusammenhang gestanden haben musste.

Dass die Ausbildung einer  $\beta$ -Faltblattkonformation in cwCT(F34W) zudem vom unmittelbaren Kontakt des Proteins mit der SUV-Grenzflächenschicht abhängig war, wurde dadurch untermauert, dass sich der Anteil an  $\beta$ -Faltblattkonformation in cwCT(F34W) scheinbar proportional zur zugänglichen Oberfläche der SUV verhielt. Dies spiegelte sich auch indirekt darin wieder, dass cwCT(F34W) oberhalb der Schmelztemperatur des DSPG wieder hauptsächlich in einer intrinsisch entfalteten Konformation vorlag (Abb. 3-19 C). Dass der temperaturbedingte Übergang der SUV-Lipidmembran von einer geordnet gelartigen zu einer ungeordnet flüssigkristallinen Phase, der mit einer Änderung der physikalisch-chemischen Eigenschaften der Lipidmembran einhergeht<sup>[247,248]</sup>, dazu führte, dass keine  $\beta$ -Faltblattkonformation in cwCT(F34W) mehr nachweisbar war, zeigte deutlich, dass eine Stabilisierung der induzierten  $\beta$ -Faltblattkonformation durch permanente Interaktion des Proteins mit der SUV-Lipidmembran stattgefunden hatte und keine von den SUV unabhängige, dauerhafte Konformationsänderung vorlag.

#### 4.3.2.2 *Parallelen der SUV-vermittelten Sekundärstrukturinduktion zum partitioning-folding coupling von Membranproteinen*

Die durch Interaktion mit der Grenzflächenregion von Lipidvesikeln induzierte Ausbildung von  $\beta$ -Faltblattstruktur im zuvor intrinsisch entfalteten cwCT(F34W) gleicht den beiden simultan stattfindenden Prozessen, die auch integrale Membranproteine vor einer etwaigen Multimerisierung und Integration in biologische Membranen durchlaufen<sup>[193,243]</sup>. Das intrinsisch entfaltete Membranprotein (bzw. intrinsisch entfaltete Regionen innerhalb desselben) dringt dabei in die Grenzflächenschicht der Lipiddoppelmembran ein und faltet gleichzeitig in seine transmembrane Konformation. Aufgrund dieser

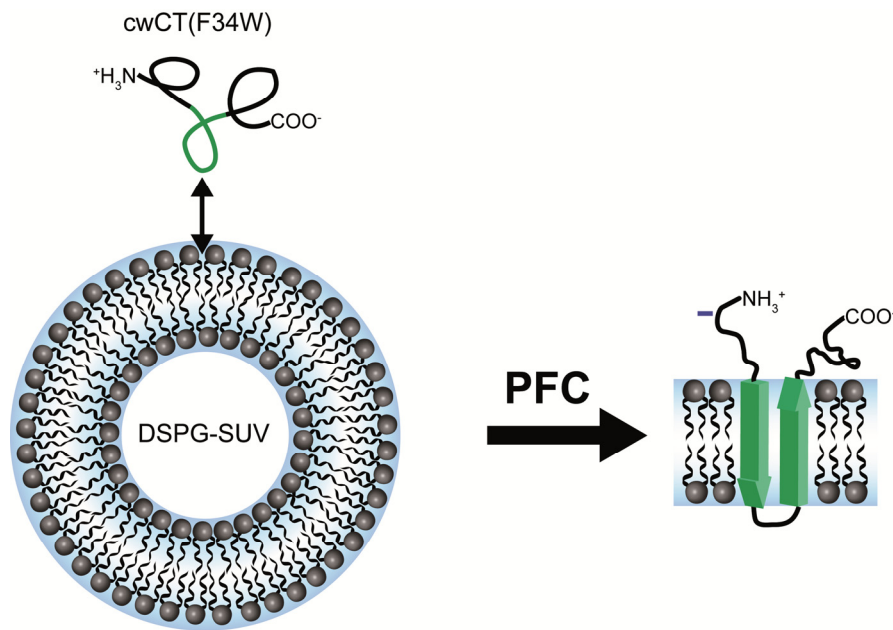
Gleichzeitigkeit werden beide Prozesse allgemein unter dem Begriff *partitioning-folding coupling* (PFC) zusammengefasst<sup>[193,243]</sup>, der indirekt noch impliziert, dass der Vorgang thermodynamisch nur möglich ist, wenn der Teil des Proteinrückgrats, welcher in die Grenzfläche eindringt, auch simultan in wasserstoffverbrückte Sekundärstrukturelemente faltet<sup>[207,243]</sup>. Nur dadurch kann die das Eindringen in die Grenzflächenschicht dominierende, aber hinsichtlich der Freien Enthalpie des Vorgangs ungünstige, Dehydratation der polaren Peptidbindung durch die mit der Sekundärstrukturausbildung einhergehende Entstehung von Wasserstoffbrückenbindungen zwischen Peptidbindungsatomen ausreichend kompensiert werden<sup>[206,243]</sup>.

Dafür, dass es sich bei dem in cwCT(F34W) durch DSPG-SUV induzierten Sekundärstrukturübergang um PFC handelte, sprechen sowohl die experimentell nachgewiesene Notwendigkeit einer direkten Interaktionsmöglichkeit von Protein und SUV-Grenzflächenschicht (vgl. Kapitel 4.3.2.1) als auch bioinformatische Analysen der cwCT(F34W)-Aminosäuresequenz (vgl. Kapitel 3.7). Das dazu verwendete Computerprogramm MPEx<sup>[151]</sup> berücksichtigt sowohl die *whole-residue* Hydrophobizität der Aminosäuresequenz, d. h. die Hydrophobizität der beteiligten Aminosäuren einschließlich der thermodynamischen Kosten, die aufgebracht werden müssen, um die polaren Peptidbindungen in eine nicht polare Umgebung bringen zu können<sup>[207]</sup>, als auch die Neigung der Sequenz zur Ausbildung von Sekundärstrukturen. Dadurch wird die sich positiv auf die thermodynamische Bilanz des PFC auswirkende Verminderung der Polarität der Peptidbindungen, die mit der Ausbildung von Wasserstoffbrückenbindungen innerhalb der sich bildenden Sekundärstrukturen einhergeht, berücksichtigt<sup>[193,206,243]</sup>. Auf diese Weise kann MPEx nur anhand der Aminosäuresequenz eines Proteins diejenigen Sequenzbereiche identifizieren, die mit hoher Wahrscheinlichkeit PFC an Lipiddoppelmembranen zeigen. Die MPEx-Analyse von cwCT(F34W) lokalisierte zwischen den Aminosäuren Gly31 und Gly91 (vgl. Kapitel 11.1.2, gelb markierte Sequenz in SUMO-cwCT(F34W)-CBD) einen hydrophoben Sequenzbereich, dessen Freie Energie für eine PFC-typische Interaktion mit der Grenzflächenschicht einer Lipiddoppelmembran prädestiniert ist, sowie einen kürzeren Bereich (Trp36–Gly67; vgl. Kapitel 11.1.2, doppelt unterstrichene Sequenz in SUMO-cwCT(F34W)-CBD) innerhalb dieser Sequenz, der infolge von PFC eine  $\beta$ -Haarnadelstruktur ausbilden kann. Die bioinformatische Analyse deckte sich mit der CD-spektroskopisch nachgewiesenen, SUV-vermittelten Induktion von

$\beta$ -Faltblattkonformation in cwCT(F34W) und unterstrich dadurch, dass es sich bei dem *in vitro* stattfindenden Prozess wahrscheinlich um einen dem PFC-ähnlichen Mechanismus handelt. Für das ebenfalls mit Hilfe von MPEx untersuchte Kontrollprotein eADF4(C16) zeigte die MPEx-Analyse interessanterweise, dass es sich, obwohl auch für eADF4(C16) in Gegenwart von DSPG-SUV experimentell ein Strukturübergang von einer intrinsisch entfalteten hin zu einer  $\beta$ -Faltblattkonformation gezeigt werden konnte (Abb. 3-16 B), mit hoher Wahrscheinlichkeit nicht um einen dem PFC-ähnlichen Mechanismus handelte. Viel eher beruhte die Induktion der  $\beta$ -Faltblattkonformation in diesem Fall auf den gleichen Mechanismen, die auch zur Ausbildung einer dauerhaften, gegenüber denaturierenden Agenzien stabilen  $\beta$ -Faltblattkonformation nach entropisch getriebener Anlagerung von eADF4(C16) an Emulsionsgrenzflächen führten<sup>[249]</sup>.

Die PFC-ähnliche Interaktion von cwCT(F34W) und den DSPG-SUV kann anhand eines Modells (Abb. 4-2) wie folgt zusammengefasst werden. Das Protein kann über einen hydrophoben Aminosäurebereich mit der gegenüber dem Lösungsmittel weniger polaren Grenzflächenschicht der SUV interagieren (Abb. 4-2, linke Seite). Dies ist möglich, da der entsprechende Sequenzbereich aufgrund seiner Aminosäurezusammensetzung auch die inhärente Eigenschaft besitzt, in eine  $\beta$ -Haarnadelstruktur zu falten (Tab. 3-2), in der die Polarität der Peptidbindung durch die zwischen den Peptidbindungsatomen entstandenen Wasserstoffbrückenbindungen herabgesetzt wurde<sup>[206,243]</sup>. Ob der entstandene  $\beta$ -Faltblattbereich dabei nur oberflächlich mit der SUV-Grenzflächenschicht interagierte oder sich komplett in die SUV-Lipiddoppelmembran integriert hatte (wie in Abb. 4-2 dargestellt), konnte jedoch weder aus den CD- noch den Fluoreszenzemissionsspektren des in cwCT(F34W) befindlichen Tryptophans geschlossen werden. Letztere zeigten, dass, verglichen mit den Fluoreszenzemissionsspektren von frei in Lösung befindlichem cwCT(F34W) in Gegenwart der DSPG-SUV eine hypsochrome Verschiebung des Emissionsmaximums um  $\sim 9$  nm und somit ein Übergang der Spektralform III zur Spektralform II stattgefunden hatte (Abb. 3-17). Die Spektralform II ist charakteristisch für Tryptophanreste, die sich in Kontakt mit Wassermolekülen sowie anderen polaren Gruppen – aber nicht mehr hauptsächlich in Kontakt zu freien Wassermolekülen – befinden<sup>[189]</sup>. Dies war zwar ein Beleg dafür, dass sich das Tryptophan nach Interaktion des Proteins mit den SUV im Bereich der polaren SUV-Grenzflächenschicht aufhielt, ließ jedoch aufgrund der Tatsache, dass das Tryptophan die erste Aminosäure des für die Bildung der  $\beta$ -

Haarnadelkonformation in Frage kommenden Bereichs darstellt (vgl. 10.1.2, gelb markierter Sequenzbereich von SUMO-cwCT(F34W)-CBD), keine Folgerungen zu, ob sich die  $\beta$ -Haarnadelstruktur komplett in die SUV-Doppelmembran integriert hatte oder nur auf dieser auflag (Abb. 4-2, rechte Seite).



**Abb. 4-2. Modell zum PFC-ähnlichen Verhalten von cwCT(F34W) in Gegenwart von DSPG-SUV.** Die DSPG-SUV bestehen aus einer Lipiddoppelmembran, bei der die negativ geladenen Kopfgruppen jeweils dem polaren Lösungsmittel zugewandt sind. Der Bereich um die Kopfgruppen ist durch eine starke Abnahme der Polarität hin zum durch die aliphatischen Kohlenwasserstoffketten gebildeten Kern gekennzeichnet (türkis: Umgebung stärkster Polarität; weiß: Umgebung schwächster Polarität) und entspricht der physikalisch-chemisch komplexen Grenzflächenschicht. Ein Teil der cwCT(F34W)-Sequenz (grün) kann mit dieser interagieren und in sie eindringen, da der Sequenzbereich gleichzeitig auch die inhärente Fähigkeit zur Ausbildung einer  $\beta$ -Haarnadelstruktur besitzt. Die amino- und carboxyterminal davon liegenden Bereiche werden durch negative Ladungen, eine nahe des Aminoterminus, sowie die Carboxylgruppe des Carboxyterminus selbst, von der ebenfalls negativ geladenen Oberfläche des SUV abgestoßen.

Für eine Integration spricht, dass der Aminosäurebereich von cwCT(F34W), der während des PFC zur Ausbildung der  $\beta$ -Haarnadelstruktur in Frage kommt, einige Eigenschaften gut beschriebener transmembraner  $\beta$ -Haarnadelstrukturen aufweist. Einerseits liegt seine Ausdehnung mit 32 Aminosäuren in der Größenordnung transmembraner  $\beta$ -Haarnadelstrukturen, die i. d. R. aus zwei je  $10 \pm 1$  Aminosäuren langen und durch eine Schleife von maximal sieben Aminosäuren miteinander verbundenen  $\beta$ -Strängen bestehen<sup>[250,251]</sup>. Zum anderen findet sich in den erwähnten Transmembranproteinen, wie

auch in cwCT(F34W), häufig eine aromatische Aminosäure am Beginn des transmembranen Sequenzbereichs wieder, die nach Insertion der  $\beta$ -Haarnadel weiterhin im Bereich der Lipidkopfguppen verbleibt<sup>[193,252]</sup>. Ein weiteres Indiz, welches für eine Integration spricht, ist die Tatsache, dass cwCT(F34W) nur in Gegenwart von DSPG-SUV, aber nicht in Gegenwart von DMPG- oder DLPG-SUV, ein PFC-ähnliches Verhalten zeigte. Da sich genannte SUV nur in der Länge der aliphatischen Seitenketten ihrer Phospholipide unterscheiden, unterscheiden sie sich (entlang der Membrannormalen) primär auch nur in den physikalisch-chemischen Eigenschaften des hydrophoben Membrankerns. Da letztere jedoch einen Einfluss darauf hatten, ob es zu einer Sekundärstrukturinduktion in cwCT(F34W) kam, muss von der Interaktionsmöglichkeit des Proteins auch mit dem Membrankern ausgegangen werden.

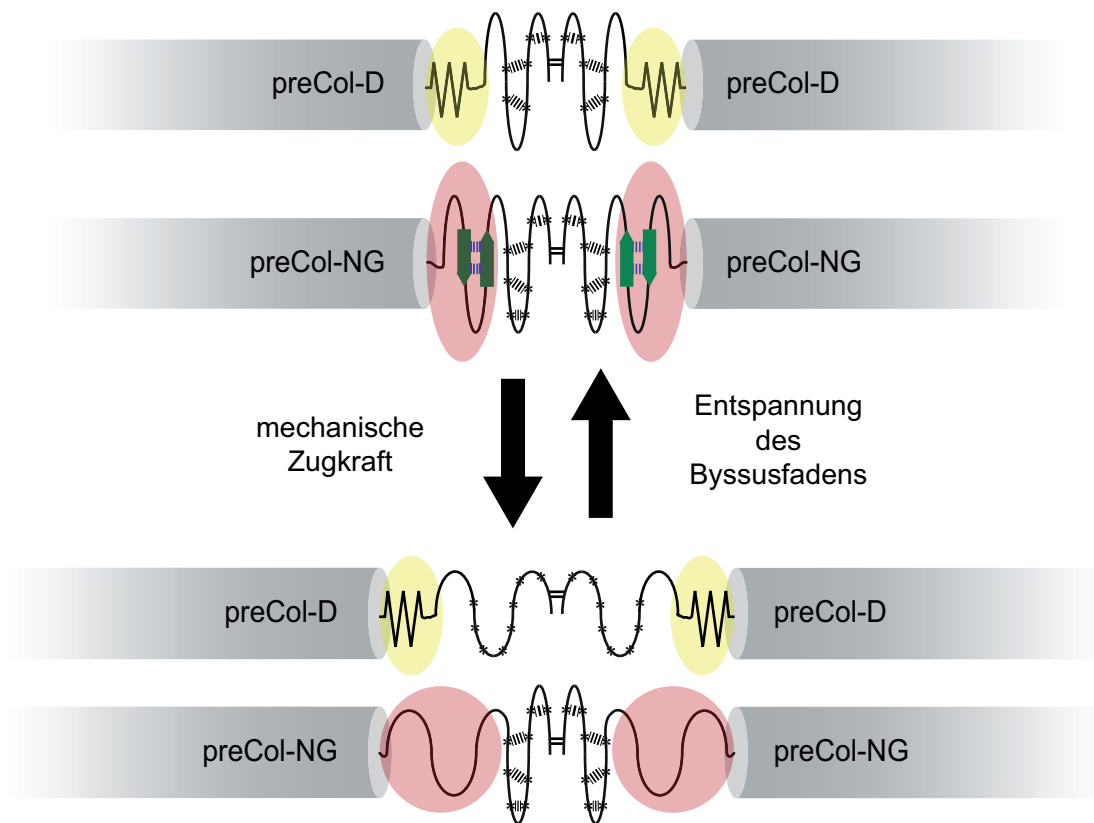
#### **4.4 Bedeutung des PFC-Verhaltens der carboxyterminal flankierenden Region für die Funktion von preCol-NG**

Das *in vitro* gezeigte PFC-Verhalten von cwCT(F34W) legt nahe, dass es *in vivo* ebenfalls zu Interaktionen der flankierenden Region von preCol-NG mit Lipidmembranen kommt. Diese können sowohl während der Assemblierung von preCol-NG bei der Entstehung der Byssusfäden als auch während Selbstheilungsprozessen innerhalb der Byssusfäden stattfinden, und ein dadurch verursachtes PFC eine wichtige Rolle beim jeweiligen Vorgang übernehmen.

Zur Neubildung von Byssusfäden werden die preCols sowie andere beteiligte Proteine nach ihrer Bildung durch spezialisiertes Drüsengewebe holokrin in die ventrale Rinne des Muschelfußes sekretiert, wo sie innerhalb von Minuten zum fertigen Faden assemblieren<sup>[41,43]</sup>. Die Sekretion der preCols erfolgt über sog. sekretorische Granulae (SG), bei denen es sich um große, nicht-sphärische Vesikel handelt, die sich im Verlauf des Sekretionsprozesses höchstwahrscheinlich aus kleineren Vesikeln durch Kondensation gebildet haben<sup>[41,43,253]</sup>. Innerhalb der SG zeigen die Proteine bereits eine anisotrope Ausrichtung und Vorstrukturierung zu Mikrofilamenten sowie auch zu stark verwundenen (*supercoiled*) Filamenten<sup>[205,253,254]</sup>, die durch Anlagerung hexagonaler 6+1 Bündel der preCol-Tripelhelices entstehen<sup>[31,57,126]</sup>. Die Quartärstruktur der preCol-Trimere ähnelt somit derjenigen fibrillärer Vertebratenkollagene<sup>[255,256]</sup>. Im Gegensatz zu diesen ist über

den Assemblierungsmechanismus der preCols, der bereits innerhalb der SG zur nachweisbaren Ausrichtung und Orientierung der Moleküle führt, noch nichts bekannt<sup>[31]</sup>. Eine Aminosäuresequenz, ähnlich den Propeptidsequenzen bei Vertretern fibrillärer Kollagene, die für die Initiierung und Steuerung der Assemblierung der Kollagentripelhelix außerordentlich wichtig sind, konnte in keinem der drei preCol-Moleküle nachgewiesen werden<sup>[31]</sup>. Da für das der carboxyterminal flankierenden Domäne von preCol-NG entsprechende Protein cwCT(F34W) im Rahmen dieser Arbeit eine spezifische Interaktion mit Lipidvesikeln nachgewiesen werden konnte, wäre es für preCol-NG denkbar, dass die räumliche Anordnung und Ausrichtung der Moleküle durch die auf dem PFC-ähnlichen Verhalten beruhende Verankerung derselben in der SG-Membran zustande kommt und so zur Initiation der Tripelhelixassemblierung (zwischen den Kollagendomänen) beiträgt. Die infolge des PFC ausgebildete  $\beta$ -Haarnadelstruktur würde die Ausbildung der Tripelhelix gleichzeitig erleichtern, da die Molekültermini dadurch – ähnlich wie bei den Seidenproteinen der Arthropoden – dichter als amorphe Domänen gepackt werden können<sup>[5]</sup>. Eine durch PFC induzierte  $\beta$ -Haarnadelkonformation innerhalb der flankierenden Region von preCol-NG kann zudem mit dem Modell, welches die Selbstheilung distaler Byssusfäden beschreibt<sup>[94]</sup>, in Einklang gebracht werden (Abb. 4-3). In diesem Bereich der Byssusfäden existieren hexagonale 6+1 Bündel<sup>[59,126]</sup> des preCol-D und des preCol-NG nebeneinander. Externe Zugkräfte, die auf diesen Fadenteil wirken, werden innerhalb der preCol-D Moleküle über die steife Kollagendomäne sowie die  $\beta$ -faltblattreichen flankierenden Regionen<sup>[57,73]</sup> direkt auf die His-Metall-Koordinationskomplexe innerhalb der His/Dopa-Domänen übertragen<sup>[94,111,257]</sup>. Jenseits der Streckgrenze brechen diese schließlich auf und dissipieren so einen Teil der auf den distalen Byssus wirkenden Energie<sup>[94]</sup>. Im Fall der preCol-NG Moleküle bleiben die His-Metall-Koordinationskomplexe hingegen erhalten, da die (im Vergleich zu preCol-D) weniger steifen flankierenden Regionen vorher mechanisch entfalten<sup>[94]</sup> (Abb. 4-3). Die bereits zu Beginn der Fadenassemblierung ausgebildete  $\beta$ -Haarnadelstruktur (s. o.) könnte aufgrund ihrer im Vergleich zu den  $\beta$ -faltblattverstärkten flankierenden Regionen von preCol-D geringeren mechanischen Stabilität (durch die geringere Anzahl an Wasserstoffbrückenbindungen bedingt) zeitgleich mit dem Aufbrechen der His-Metall-Koordinationskomplexe in preCol-D unter Freilegung zusätzlicher molekularer Konturlänge entfalten (Abb. 4-3, unterer Teil).





**Abb. 4-3. Modell zur Rolle einer PFC-induzierten  $\beta$ -Haarnadelstruktur innerhalb der flankierenden Regionen von preCol-NG während des Selbstheilungsprozesses distaler Byssusfäden.** Im distalen Teil der Byssusfäden liegen hexagonale 6+1 Bündel von preCol-D und -NG nebeneinander vor. Beim Einwirken externer Zugkräfte wird die mechanische Energie in preCol-D Molekülen über die steife Kollagendomäne (grauer Zylinder) sowie die  $\beta$ -faltblattreichen flankierenden Domänen (gelb markiert) direkt auf die in der His/DOPA-Region vorhandenen His-Metall-Koordinationskomplexe (schwarze gestrichelte Linien) übertragen, während die Energie in preCol-NG Molekülen über die Kollagendomäne direkt auf die weniger steifen flankierenden Domänen (rot markiert) einwirkt. Während letztere unter Aufbrechen der Wasserstoffbrückenbindung innerhalb der  $\beta$ -Haarnadelstruktur entfalten, bleiben die His-Metall-Koordinationskomplexe, anders als in den preCol-D Molekülen, erhalten. In der Entspannungsphase des Byssus begünstigen im Faden enthaltene Lipide das Rückfalten der  $\beta$ -Haarnadelstruktur, was zusammen mit entropisch bedingten Vorgängen ein elastisches „Zusammenziehen“ des Fadens bewirkt. Da die preCol-Bündel auch lateral miteinander interagieren, werden auch die preCol-D Moleküle auf diese Weise in ihre ursprüngliche Position zurückgebracht. Schließlich können sich die His-Metall-Koordinationskomplexe neu ausbilden, was (nahezu) zur Wiederherstellung der ursprünglichen mechanischen Eigenschaften des distalen Byssus führt.

Nachdem ein Teil der auf den distalen Byssus wirkenden Kräfte auf diese Weise dissipiert und ein katastrophales Versagen des Byssus verhindert wurde, kann nach Wegfall der externen Zugkräfte der Selbstheilungsprozess einsetzen. Im Byssusfaden vorhandene Lipide<sup>[37,123]</sup> könnten anschließend als Trigger der Rückfaltung der  $\beta$ -Haarnadelstruktur fungieren. Gemeinsam mit der durch die entfalteten, gestreckten

hydrophoben Bereiche erzeugten entropischen Rückstellkraft trägt dies zu einer schnellen Repositionierung der preCol-NG Moleküle bei<sup>[94]</sup>. Da die preCol-Bündel auch lateral miteinander interagieren<sup>[126]</sup> werden die preCol-D Moleküle wieder in ihre ursprüngliche Position zurückgebracht. Dies erlaubt die Neubildung der His-Metall-Koordinationskomplexe in den His/Dopa-Domänen von preCol-D und die annähernde Wiederherstellung der ursprünglichen mechanischen Eigenschaften des distalen Teils der Byssusfäden<sup>[94,126]</sup>. Dass die Bereitstellung zusätzlicher molekularer Konturlänge durch die Entfaltung einer  $\beta$ -Haarnadelstruktur in der flankierenden Region von preCol-NG dabei gut in ein Modell zur elastischen Dissipation mechanischer Energie passt, zeigt sowohl die Tatsache, dass vergleichbare Prozesse auch bei anderen ebenfalls mechanischen Belastungen ausgesetzten Strukturproteinen auftreten<sup>[258,259]</sup>, als auch die Tatsache, dass eine elastische Entfaltung für die flankierenden Regionen der preCols schon häufig postuliert wurde<sup>[59,94,108,111,126]</sup>, und die Bereitstellung zusätzlicher molekularer Konturlänge zudem allgemein eine verbreitete biologische Strategie zur Erhöhung der Materialzähigkeit darzustellen scheint<sup>[126]</sup>.

Das PFC-ähnliche Verhalten von cwCT(F34W) zeigt nicht nur, dass der flankierenden Region von preCol-NG *in vivo* eine wichtige Rolle sowohl bei der Assemblierung von Byssusfäden als auch während deren Selbstheilung zukommen könnte, sondern eröffnet ebenso die Möglichkeit, cwCT(F34W) (bzw. dessen Sequenz) zur (Weiter-)Entwicklung selbstheilender Hydrogele zu nutzen. Die aus den preCol-Molekülen bekannten nicht-kovalenten His-Metall-Quervernetzungen fließen bereits in die Entwicklung solcher muschelinspirierter Hydrogele<sup>[260-262]</sup>, für die teilweise auch Selbstheilungseigenschaften gezeigt werden konnten<sup>[261]</sup>, ein. Durch gezielte Funktionalisierung der zugrunde liegenden organischen Polymere mit cwCT(F34W) (bzw. die gezielte Inkorporation der cwCT(F34W)-Sequenz im Falle proteinbasierter Hydrogele) sowie dem Einbau von Lipiden könnte, ähnlich dem für die flankierende Region von preCol-NG postulierten Mechanismus, eine „molekulare Feder“ erzeugt werden, die einerseits mechanische Kräfte dissipiert und andererseits nach Aufbrechen der His-Metall-Quervernetzung ein schnelleres Rückführen der Moleküle in ihre ursprüngliche Position triggert, was für die langsame verlaufende, korrekte Neubildung der Quervernetzungen<sup>[94,126]</sup> und somit für die Wiederherstellung nahezu ursprünglicher mechanischer Eigenschaften des Polymermaterials von großer Bedeutung wäre.

## 5 Zusammenfassung

Muscheln der Gattung *Mytilus* haben zur Besiedlung von Gezeitenzonen ein als Byssus bezeichnetes Verankerungssystem entwickelt, welches unter den dort vorherrschenden Bedingungen ein stabiles Anhaften am Untergrund und Schutz vor durch Wellen und Strömungen verursachten Kräften gewährleistet. Der strukturelle Kern eines jeden Byssusfadens, der für die Aufnahme und Dissipation mechanischer Kräfte hauptverantwortlich ist, wird von drei unterscheidbaren, als preCols bezeichneten Proteinen gebildet. Diese zeigen auf molekularer Ebene eine modulare Domänenstruktur. Die ungefähr die Hälfte des Proteins dominierende Kollagendomäne, an die carboxyterminal ein Cluster saurer Aminosäuren anschließt, sowie Termini, die durch einen hohen Anteil von Histidin sowie 3,4-Dihydroxyphenylalanin (DOPA) gekennzeichnet sind, sind allen drei preCols gemein. Unterschiede zeigen sich hauptsächlich in den die Kollagendomäne und den Cluster saurer Aminosäuren flankierenden Regionen, da diese entweder aus elastinähnlichen, aus seidenähnlichen oder aus glycinreichen Konsensusmotiven bestehen. Die Unterschiede zwischen den flankierenden Regionen sowie das spezifische Verteilungsmuster der preCols entlang der Fadenachse spiegeln sich direkt in den entlang der Fadenachse variierenden mechanischen Eigenschaften der Byssusfäden wieder. Deshalb wird den flankierenden Regionen auch ein Beitrag zu diesen mechanischen Eigenschaften zugeschrieben.

Ziel der hier vorliegenden Arbeit war es, die flankierenden Regionen der *M. galloprovincialis* preCols rekombinant in *E. coli* herzustellen, um durch deren biochemische und biophysikalische Untersuchung die Struktur-Funktions-Beziehungen detailliert aufzuklären und dadurch Informationen über Rolle und Funktion der flankierenden Regionen *in vivo* zu erhalten.

Zunächst konnte eine Expressions- und Reinigungsstrategie entwickelt werden, die es ausgehend von einer cDNA-Datenbank des Muschelfußgewebes von *M. galloprovincialis* erlaubt, auf der Aminosäuresequenz der flankierenden Regionen beruhende *flank*-Proteine rekombinant herzustellen. Durch die Verwendung sowohl nativer als auch

optimierter Nukleotidsequenzen konnte gezeigt werden, dass verkürzte Zielproteine während der Proteinbiosynthese in erster Linie auf *in vivo* Abbauprozesse und weniger auf eine ineffiziente Translation zurückzuführen waren. Durch simultane Verwendung zweier großer, strukturierter *tags* konnte durch Abschirmung der eigentlichen *flank*-Sequenz deren *in vivo*-Abbau minimiert und über sequentielle *tag*-Abspaltungen und säulenchromatographische Reinigungsschritte eine *flank*-Domäne rekombinant hergestellt werden.

Diese sequenzoptimierte carboxyterminal flankierende Region von preCol-NG (cwCT(F34W)) zeigte in verschiedenen wässrigen Lösungen eine intrinsisch entfaltete Struktur mit möglicherweise untergeordneten  $3_1$ -helikalen Strukturanteilen und war zeitlich und thermisch stabil in Lösung. In Gegenwart bestimmter kleiner unilamellarer Lipidvesikel (SUV) kam es zur Induktion von Sekundärstruktur mit einer  $\beta$ -faltblattdominierten Konformation. Die Art der Strukturinduktion ähnelte dem *partitioning-folding coupling* (PFC) Mechanismus der Faltung von Membranproteinen. Mit Hilfe einer *whole-residue* Hydrophobizitätsanalyse gelang die Identifizierung eines Aminosäurebereichs innerhalb von cwCT(F34W), der mit einer SUV-Lipiddoppelmembran interagieren kann und der zur Ausbildung einer  $\beta$ -Haarnadelkonformation neigt. Zudem sprachen experimentelle und bioinformatische Daten dafür, dass es zu einer Integration der  $\beta$ -Haarnadelstruktur in die SUV-Lipiddoppelmembran kommt.

Die anhand von cwCT(F34W) gewonnenen Erkenntnisse legen nicht nur nahe, dass die carboxyterminal flankierende Region von preCol-NG während der Proteinassemblierung und Entstehung der Byssusfäden sowie deren Selbstheilung eine wichtige Rolle spielen kann, sondern eröffnen auch Möglichkeiten für die (Weiter-)Entwicklung potentiell selbstheilender Biomaterialien.

## 6 Summary

Mussels of the genus *Mytilus* developed a holdfast structure called the byssus which enables them to tether to different substrates in intertidal zones, thereby withstanding the mechanical forces caused by water current and swell. The load-bearing and energy dissipating core of the proteinaceous threads forming the byssus primarily comprises three distinct proteins termed preCols. On a molecular level, these proteins reveal a modular structure showing distinguishable domains. Nearly half of the preCol sequence is a collagen domain, carboxyterminally followed by a cluster of acidic amino acids. Further, preCols share similar terminal domains characterized by a high content in histidine and 3,4-dihydroxy-phenylalanine (DOPA) residues. In turn, differences between the different preCol types are mainly localized within the regions directly flanking the collagen domain, as those are composed of silk-like, elastin-like, or glycine-rich consensus motifs. Sequence differences between the flanking regions as well as the distinct distribution pattern of preCols along the thread axis perfectly correlate with the threads' longitudinally changing mechanical properties. Therefore, it is widely accepted that one main function of the flanking regions is to determine and maintain byssus mechanics.

To elucidate the structure-function-relationships of the flanking regions in more detail and to gain more information about their roles *in vivo*, this work aimed to recombinantly produce proteins resembling the flanking regions of *M. galloprovincialis* preCols as well as their subsequent biochemical and biophysical analysis.

Based on a *M. galloprovincialis* derived cDNA library of mussel foot tissue, a strategy for the expression and purification of recombinant proteins, that mimic the preCol flanking regions, was successfully developed. By employing native as well as codon usage optimized nucleotide sequences, it could be shown that the appearance of truncated versions of the target proteins during heterologous expression was mainly caused by *in vivo* degradation rather than by inefficient translation processes. In presence of two large, globular tags, *in vivo* degradation could be minimized, which, in turn, allowed isolation

of the flank protein by a subsequent series of tag-cleavage and column purification steps.

In aqueous solutions, the carboxyterminal flanking region of preCol-NG (cwCT(F34W)) adopted an intrinsically disordered conformation with some  $3_1$ -helical content and remained stable in solution, even at high temperatures. However, secondary structure formation was induced in presence of certain small unilamellar lipid vesicles (SUV), yielding a  $\beta$ -sheet dominated conformation in a process resembling a step during membrane protein folding known as partitioning-folding coupling (PFC). Whole-residue hydrophobicity analysis was used to identify an amino acid stretch within cwCT(F34W) able to interact with the SUVs' lipid double layer and prone to fold into a  $\beta$ -hairpin conformation upon interaction. Experimental data and bioinformatics further suggested an integration of the formed  $\beta$ -hairpin into the lipid double layer.

The abovementioned findings not only indicate that the carboxyterminal flanking region of preCol-NG can play an important role during protein assembly, byssus formation and byssal self-healing, but also enables a (future) development of self-healing biomaterials.

## 7 Literatur

- [1] Coddington, J. A. & Levi, H. W. (1991) Systematics and Evolution of Spiders (Araneae). *Annu. Rev. Ecol. Syst.* **22**, 565-592.
- [2] Craig, C. L. (1997) Evolution of arthropod silks. *Annu. Rev. Entomol.* **42**, 231-267.
- [3] Sehnael, F. & Sutherland, T. (2008) Silks produced by insect labial glands. *Prion* **2**, 145-153.
- [4] Sutherland, T. D., Weisman, S., Walker, A. A. & Mudie, S. T. (2012) The coiled coil silk of bees, ants, and hornets. *Biopolymers* **97**, 446-454.
- [5] Sutherland, T. D., Young, J. H., Weisman, S., Hayashi, C. Y. & Merritt, D. J. (2010) Insect silk: one name, many materials. *Annu. Rev. Entomol.* **55**, 171-188.
- [6] Ayoub, N. A., Garb, J. E., Tinghitella, R. M., Collin, M. A. & Hayashi, C. Y. (2007) Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. *PLoS ONE* **2**, e514.
- [7] Garb, J. E., Ayoub, N. A. & Hayashi, C. Y. (2010) Untangling spider silk evolution with spidroin terminal domains. *BMC Evol. Biol.* **10**, 243.
- [8] Sponner, A., Unger, E., Grosse, F. & Weisshart, K. (2004) Conserved C-termini of spidroins are secreted by the major ampullate glands and retained in the silk thread. *Biomacromolecules* **5**, 840-845.
- [9] Xu, M. & Lewis, R. V. (1990) Structure of a protein superfiber: spider dragline silk. *Proc. Natl. Acad. Sci. USA* **87**, 7120-7124.
- [10] Jin, H. J. & Kaplan, D. L. (2003) Mechanism of silk processing in insects and spiders. *Nature* **424**, 1057-1061.
- [11] Vollrath, F. & Knight, D. P. (2001) Liquid crystalline spinning of spider silk. *Nature* **410**, 541-548.
- [12] Holland, C., Terry, A. E., Porter, D. & Vollrath, F. (2006) Comparing the rheology of native spider and silkworm spinning dope. *Nat. Mater.* **5**, 870-874.
- [13] Kojic, N., Bico, J., Clasen, C. & McKinley, G. H. (2006) Ex-vivo rheology of spider silk. *J. Exp. Biol.* **209**, 4355-4362.
- [14] Hayashi, C. Y. & Lewis, R. V. (2001) Spider flagelliform silk: lessons in protein design, gene structure, and molecular evolution. *Bioessays* **23**, 750-756.
- [15] Hepburn, H. R. & Kurstjens, S. P. (1988) The Combs of Honeybees as Composite-Materials. *Apidologie* **19**, 25-36.
- [16] Zhang, K., Duan, H., Karihaloo, B. L. & Wang, J. (2010) Hierarchical, multilayered cell walls reinforced by recycled silk cocoons enhance the structural integrity of honeybee combs. *Proc. Natl. Acad. Sci. USA* **107**, 9502-9506.
- [17] Ishay, J. & Ruttner, F. (1971) Thermoregulation in Nests of Hornets. *Z. Vergl. Physiol.* **72**, 423-434.
- [18] Plotkin, M., Ermakov, N. Y., Volynchik, S., Bergman, D. J. & Ishay, J. S. (2007) Prevention of hyperthermia with silk of the oriental hornet, *Vespa orientalis*: a hypothesis. *Microsc. Res. Tech.* **70**, 69-75.
- [19] Ellis, M. B., Nicolson, S. W., Crewe, R. M. & Dietemann, V. (2010) Brood comb as a humidity buffer in honeybee nests. *Naturwissenschaften* **97**, 429-433.
- [20] Humenik, M., Scheibel, T. & Smith, A. Spider Silk: Understanding the Structure-Function Relationship of a Natural Fiber in *Molecular Assembly in Natural and Engineered Systems*, Vol. 103 (Editor Howorka, S.), Academic Press (2011), 131-185.
- [21] Vollrath, F. (1999) Biology of spider silk. *Int. J. Biol. Macromol.* **24**, 81-88.

- [22] Vollrath, F. (2000) Strength and structure of spiders' silks. *J. Biotechnol.* **74**, 67-83.
- [23] Gosline, J. M., Guerette, P. A., Ortlepp, C. S. & Savage, K. N. (1999) The mechanical design of spider silks: From fibroin sequence to mechanical function. *J. Exp. Biol.* **202**, 3295-3303.
- [24] Gerritsen, V. B. (2002) The tiptoe of an airbus. *Protein Spotlight* **24**, 1-2.
- [25] Heim, M., Keerl, D. & Scheibel, T. (2009) Spider Silk: From Soluble Protein to Extraordinary Fiber. *Angew. Chem. Int. Ed. Engl.* **48**, 3584-3596.
- [26] Hayashi, C. Y. Evolution of spider silk proteins: insight from phylogenetic analyses in *Molecular Systematics and Evolution: Theory and Practice* (Editoren DeSalle, R., Giribet, G. & Wheeler, W.), Birkhäuser (2002), 209-223.
- [27] Hayashi, C. Y. & Lewis, R. V. (1998) Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. *J. Mol. Biol.* **275**, 773-784.
- [28] Qin, X. X. & Waite, J. H. (1995) Exotic Collagen Gradients in the Byssus of the Mussel *Mytilus-Edulis*. *J. Exp. Biol.* **198**, 633-644.
- [29] Carrington, E. (2008) Along the silk road, spiders make way for mussels. *Trends Biotechnol.* **26**, 55-57.
- [30] Benedict, C. V. & Waite, J. H. (1986) Location and analysis of byssal structural proteins of *Mytilus edulis*. *J. Morphol.* **189**, 171-181.
- [31] Waite, J. H., Qin, X. X. & Coyne, K. J. (1998) The peculiar collagens of mussel byssus. *Matrix Biol.* **17**, 93-106.
- [32] Claussen, K. U., Giesa, R., Scheibel, T. & Schmidt, H.-W. (2012) Learning from nature: synthesis and characterization of longitudinal polymer gradient materials inspired by mussel byssus threads. *Macromol. Rapid Commun.* **33**, 206-211.
- [33] Claussen, K. U., Lintz, E. S., Giesa, R., Schmidt, H.-W. & Scheibel, T. (2013) Protein gradient films of fibroin and gelatine. *Macromol. Biosci.* **13**, 1396-1403.
- [34] Claussen, K. U., Scheibel, T., Schmidt, H.-W. & Giesa, R. (2012) Polymer Gradient Materials: Can Nature Teach Us New Tricks? *Macromol. Mater. Eng.* **297**, 938-957.
- [35] Deming, T. J. (1999) Mussel byssus and biomolecular materials. *Curr. Opin. Chem. Biol.* **3**, 100-105.
- [36] Hagenau, A. & Scheibel, T. (2010) Towards the Recombinant Production of Mussel Byssal Collagens. *J. Adhes.* **86**, 10-24.
- [37] Holten-Andersen, N. & Waite, J. H. (2008) Mussel-designed Protective Coatings for Compliant Substrates. *J. Dent. Res.* **87**, 701-709.
- [38] Kamino, K. (2008) Underwater adhesive of marine organisms as the vital link between biological science and material science. *Mar. Biotechnol. (NY)* **10**, 111-121.
- [39] Lee, B. P., Messersmith, P. B., Israelachvili, J. N. & Waite, J. H. (2011) Mussel-inspired adhesives and coatings. *Annu. Rev. Mater. Res.* **41**, 99-132.
- [40] Yamamoto, H. (1996) Marine adhesive proteins and some biotechnological applications. *Biotechnol. Genet. Eng. Rev.* **13**, 133-165.
- [41] Waite, J. H. The formation of mussel byssus: Anatomy of a natural manufacturing process in *Structure, cellular synthesis and assembly of biopolymers. Results and problems in cell differentiation*, Vol. 19 (Editor Case, S. T.), Springer (1992), 27-54.
- [42] Silverman, H. G. & Roberto, F. F. (2007) Understanding marine mussel adhesion. *Mar. Biotechnol.* **9**, 661-681.
- [43] Silverman, H. G. & Roberto, F. F. Byssus Formation in *Mytilus* in *Biological Adhesive Systems: From Nature to Technical and Medical Application* (Editoren Von Byern, J. & Grunwald, I.), Springer (2010), 273-283.



- 
- [44] Tanaka, K. *et al.* (1999) Determination of the site of disulfide linkage between heavy and light chains of silk fibroin produced by *Bombyx mori*. *Biochim. Biophys. Acta* **1432**, 92-103.
  - [45] Inoue, S. *et al.* (2000) Silk fibroin of *Bombyx mori* is secreted, assembling a high molecular mass elementary unit consisting of H-chain, L-chain, and P25, with a 6:6:1 molar ratio. *J. Biol. Chem.* **275**, 40517-40528.
  - [46] Kundu, S. C., Dash, B. C., Dash, R. & Kaplan, D. L. (2008) Natural protective glue protein, sericin bioengineered by silkworms: Potential for biomedical and biotechnological applications. *Prog. Polym. Sci.* **33**, 998-1012.
  - [47] Engster, M. S. (1976) Studies on silk secretion in the Trichoptera (F. Limmephilidae). II. Structure and amino acid composition of the silk. *Cell Tiss. Res.* **169**, 77-92.
  - [48] Yonemura, N., Sehna, F., Mita, K. & Tamura, T. (2006) Protein composition of silk filaments spun under water by caddisfly larvae. *Biomacromolecules* **7**, 3370-3378.
  - [49] Atkins, E. D. T. (1967) A Four-strand Coiled-coil Model for Some Insect Fibrous Proteins. *J. Mol. Biol.* **24**, 139-141.
  - [50] Hinman, M. B. & Lewis, R. V. (1992) Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. *J. Biol. Chem.* **267**, 19320-19324.
  - [51] Sponner, A. *et al.* (2007) Composition and hierarchical organisation of a spider silk. *PLoS ONE* **2**, e998.
  - [52] Brooks, A. E., Steinkraus, H. B., Nelson, S. R. & Lewis, R. V. (2005) An investigation of the divergence of major ampullate silk fibers from *Nephila clavipes* and *Argiope aurantia*. *Biomacromolecules* **6**, 3095-3099.
  - [53] Colgin, M. A. & Lewis, R. V. (1998) Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like "spacer regions". *Protein Sci.* **7**, 667-672.
  - [54] Zschokke, S. (1993) The influence of the auxiliary spiral on the capture spiral in *Araneus diadematus* Clerck (Araneidae). *Bull. Br. Arachnol. Soc.* **9**, 167-173.
  - [55] Schulz, S. (2001) Composition of the silk lipids of the spider *Nephila clavipes*. *Lipids* **36**, 637-647.
  - [56] Coyne, K. J., Qin, X. X. & Waite, J. H. (1997) Extensible collagen in mussel byssus: A natural block copolymer. *Science* **277**, 1830-1832.
  - [57] Qin, X. X., Coyne, K. J. & Waite, J. H. (1997) Tough tendons. Mussel byssus has collagen with silk-like domains. *J. Biol. Chem.* **272**, 32623-32627.
  - [58] Qin, X. X. & Waite, J. H. (1998) A potential mediator of collagenous block copolymer gradients in mussel byssal threads. *Proc. Natl. Acad. Sci. USA* **95**, 10517-10522.
  - [59] Hassenkam, T., Gutschmann, T., Hansma, P., Sagert, J. & Waite, J. H. (2004) Giant bent-core mesogens in the thread forming process of marine mussels. *Biomacromolecules* **5**, 1351-1355.
  - [60] Sagert, J. & Waite, J. H. (2009) Hyperunstable matrix proteins in the byssus of *Mytilus galloprovincialis*. *J. Exp. Biol.* **212**, 2224-2236.
  - [61] Holten-Andersen, N., Zhao, H. & Waite, J. H. (2009) Stiff coatings on compliant biofibers: the cuticle of *Mytilus californianus* byssal threads. *Biochemistry-US* **48**, 2752-2759.
  - [62] Holten-Andersen, N., Fantner, G. E., Hohlbauch, S., Waite, J. H. & Zok, F. W. (2007) Protective coatings on extensible biofibres. *Nature Mater.* **6**, 669-672.
  - [63] Bini, E., Knight, D. P. & Kaplan, D. L. (2004) Mapping domain structures in silks from insects and spiders related to protein assembly. *J. Mol. Biol.* **335**, 27-40.
  - [64] Eisoldt, L., Smith, A. & Scheibel, T. (2011) Decoding the secrets of spider silk. *Mater. Today* **14**, 80-86.
  - [65] Rising, A. *et al.* (2005) Spider silk proteins - Mechanical property and gene sequence. *Zoolog. Sci.* **22**, 273-281.

- [66] Du, N. *et al.* (2006) Design of superior spider silk: From nanostructure to mechanical properties. *Biophys. J.* **91**, 4528-4535.
- [67] Grubb, D. T. & Jelinski, L. W. (1997) Fiber morphology of spider silk: The effects of tensile deformation. *Macromolecules* **30**, 2860-2867.
- [68] Kummerlen, J., vanBeek, J. D., Vollrath, F. & Meier, B. H. (1996) Local structure in spider dragline silk investigated by two-dimensional spin-diffusion nuclear magnetic resonance. *Macromolecules* **29**, 2920-2928.
- [69] Riekel, C., Muller, M. & Vollrath, F. (1999) In situ X-ray diffraction during forced silking of spider silk. *Macromolecules* **32**, 4464-4466.
- [70] Simmons, A., Ray, E. & Jelinski, L. W. (1994) Solid-State  $^{13}\text{C}$  NMR of *Nephila clavipes* dragline silk establishes structure and identity of crystalline regions. *Macromolecules* **27**, 5235-5237.
- [71] Simmons, A. H., Michal, C. A. & Jelinski, L. W. (1996) Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. *Science* **271**, 84-87.
- [72] Sampath, S. *et al.* (2012) X-ray diffraction study of nanocrystalline and amorphous structure within major and minor ampullate dragline spider silks. *Soft Matter* **8**, 6713-6722.
- [73] Arnold, A. A. *et al.* (2012) Solid-State NMR Structure Determination of Whole Anchoring Threads from the Blue Mussel *Mytilus edulis*. *Biomacromolecules* **14**, 132-141.
- [74] Jenkins, J. E., Creager, M. S., Lewis, R. V., Holland, G. P. & Yarger, J. L. (2010) Quantitative correlation between the protein primary sequences and secondary structures in spider dragline silks. *Biomacromolecules* **11**, 192-200.
- [75] Lefevre, T., Rousseau, M. E. & Pezolet, M. (2007) Protein secondary structure and orientation in silk as revealed by Raman spectromicroscopy. *Biophys. J.* **92**, 2885-2895.
- [76] Shen, Y., Johnson, M. A. & Martin, D. C. (1998) Microstructural characterization of *Bombyx mori* silk fibers. *Macromolecules* **31**, 8857-8864.
- [77] Hutchinson, E. G. & Thornton, J. M. (1994) A revised set of potentials for beta-turn formation in proteins. *Protein Sci.* **3**, 2207-2216.
- [78] Ohgo, K., Kawase, T., Ashida, J. & Asakura, T. (2006) Solid-state NMR analysis of a peptide (Gly-Pro-Gly-Gly-Ala)(6)-Gly derived from a flagelliform silk sequence of *Nephila clavipes*. *Biomacromolecules* **7**, 1210-1214.
- [79] Matsushima, N., Creutz, C. E. & Kretsinger, R. H. (1990) Polyproline, beta-turn helices. Novel secondary structures proposed for the tandem repeats within rhodopsin, synaptophysin, synexin, gliadin, RNA polymerase II, hordein, and gluten. *Proteins* **7**, 125-155.
- [80] Gatesy, J., Hayashi, C., Motriuk, D., Woods, J. & Lewis, R. (2001) Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science* **291**, 2603-2605.
- [81] Sun, C. J. & Waite, J. H. (2005) Mapping chemical gradients within and along a fibrous structural tissue, mussel byssal threads. *J. Biol. Chem.* **280**, 39332-39336.
- [82] He, Y. X. *et al.* (2012) N-Terminal Domain of *Bombyx mori* Fibroin Mediates the Assembly of Silk in Response to pH Decrease. *J. Mol. Biol.* **418**, 197-207.
- [83] Zhou, C. Z. *et al.* (2001) Silk fibroin: Structural implications of a remarkable amino acid sequence. *Proteins* **44**, 119-122.
- [84] Takei, F., Kikuchi, Y., Kikuchi, A., Mizuno, S. & Shimura, K. (1987) Further Evidence for Importance of the Subunit Combination of Silk Fibroin in Its Efficient Secretion from the Posterior Silk Gland-Cells. *J. Cell. Biol.* **105**, 175-180.
- [85] Sezutsu, H. *et al.* (2007) Identification of four major hornet silk genes with a complex of alanine-rich and serine-rich sequences in *Vespa simillima xanthoptera* Cameron. *Biosci. Biotech. Bioch.* **71**, 2725-2734.

- 
- [86] Ittah, S., Michaeli, A., Goldblum, A. & Gat, U. (2007) A model for the structure of the C-terminal domain of dragline spider silk and the role of its conserved cysteine. *Biomacromolecules* **8**, 2768-2773.
  - [87] Guerette, P. A., Ginzinger, D. G., Weber, B. H. F. & Gosline, J. M. (1996) Silk properties determined by gland-specific expression of a spider fibroin gene family. *Science* **272**, 112-115.
  - [88] Motriuk-Smith, D., Smith, A., Hayashi, C. Y. & Lewis, R. V. (2005) Analysis of the conserved N-terminal domains in major ampullate spider silk proteins. *Biomacromolecules* **6**, 3152-3159.
  - [89] Rising, A., Hjalms, G., Engstrom, W. & Johansson, J. (2006) N-terminal nonrepetitive domain common to dragline, flagelliform, and cylindriform spider silk proteins. *Biomacromolecules* **7**, 3120-3124.
  - [90] Challis, R. J., Goodacre, S. L. & Hewitt, G. M. (2006) Evolution of spider silks: conservation and diversification of the C-terminus. *Insect. Mol. Biol.* **15**, 45-56.
  - [91] Askarieh, G. *et al.* (2010) Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature* **465**, 236-238.
  - [92] Hedhammar, M. *et al.* (2008) Structural properties of recombinant nonrepetitive and repetitive parts of major ampullate spidroin 1 from *Euprosthenops australis*: Implications for fiber formation. *Biochemistry-Us* **47**, 3407-3417.
  - [93] Harrington, M. J. & Waite, J. H. Short-order tendons: liquid crystal mesophases, metal-complexation and protein gradients in the externalized collagens of mussel byssal threads. in *Fibrous Proteins* (Editor Scheibel, T.), Landes Bioscience (2008), 30-45.
  - [94] Harrington, M. J. & Waite, J. H. (2007) Holdfast heroics: comparing the molecular and mechanical properties of *Mytilus californianus* byssal threads. *J. Exp. Biol.* **210**, 4307-4318.
  - [95] Waite, J. H., Vaccaro, E., Sun, C. J. & Lucas, J. M. (2002) Elastomeric gradients: a hedge against stress concentration in marine holdfasts? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **357**, 143-153.
  - [96] Sundberg, R. J. & Martin, R. B. (1974) Interactions of Histidine and Other Imidazole Derivatives with Transition-Metal Ions in Chemical and Biological-Systems. *Chem. Rev.* **74**, 471-517.
  - [97] Schmitt, L., Ludwig, M., Gaub, H. E. & Tampe, R. (2000) A metal-chelating microscopy tip as a new toolbox for single-molecule experiments by atomic force microscopy. *Biophys. J.* **78**, 3275-3285.
  - [98] Harrington, M. J. & Waite, J. H. (2008) pH-dependent locking of giant mesogens in fibers drawn from mussel byssal collagens. *Biomacromolecules* **9**, 1480-1486.
  - [99] Burzio, L. A. & Waite, J. H. (2000) Cross-linking in adhesive quinoproteins: studies with model decapeptides. *Biochemistry-Us* **39**, 11147-11153.
  - [100] McDowell, L. M., Burzio, L. A., Waite, J. H. & Schaefer, J. (1999) Rotational echo double resonance detection of cross-links formed in mussel byssus under high-flow stress. *J. Biol. Chem.* **274**, 20293-20295.
  - [101] Lefevre, T., Paquet-Mercier, F., Rioux-Dube, J.-F. & Pezolet, M. (2012) Structure of silk by raman spectromicroscopy: From the spinning glands to the fibers. *Biopolymers* **97**, 322-336.
  - [102] Lefevre, T. *et al.* (2007) Conformation of spider silk proteins in situ in the intact major ampullate gland and in solution. *Biomacromolecules* **8**, 2342-2344.
  - [103] Asakura, T. *et al.* (2001) A repeated beta-turn structure in poly(Ala-Gly) as a model for silk I of *Bombyx mori* silk fibroin studied with two-dimensional spin-diffusion NMR under off magic angle spinning and rotational echo double resonance. *J. Mol. Biol.* **306**, 291-305.
  - [104] Bella, J., Eaton, M., Brodsky, B. & Berman, H. M. (1994) Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution. *Science* **266**, 75-81.
  - [105] Aldred, N., Wills, T., Williams, D. N. & Clare, A. S. (2007) Tensile and dynamic mechanical analysis of the distal portion of mussel (*Mytilus edulis*) byssal threads. *J. R. Soc. Interface* **4**, 1159-1167.

- [106] Bell, E. C. & Gosline, J. M. (1996) Mechanical design of mussel byssus: Material yield enhances attachment strength. *J. Exp. Biol.* **199**, 1005-1017.
- [107] Brazee, S. L. & Carrington, E. (2006) Interspecific comparison of the mechanical properties of mussel byssus. *Biol. Bull.* **211**, 263-274.
- [108] Hagenau, A., Papadopoulos, P., Kremer, F. & Scheibel, T. (2011) Mussel collagen molecules with silk-like domains as load-bearing elements in distal byssal threads. *J. Struct. Biol.* **175**, 339-347.
- [109] Lucas, J. M., Vaccaro, E. & Waite, J. H. (2002) A molecular, morphometric and mechanical comparison of the structural elements of byssus from *Mytilus edulis* and *Mytilus galloprovincialis*. *J. Exp. Biol.* **205**, 1807-1817.
- [110] Smeathers, J. E. & Vincent, J. F. V. (1979) Mechanical properties of mussel byssus threads. *J. Molluscan Stud.* **45**, 219-230.
- [111] Harrington, M. J., Gupta, H. S., Fratzl, P. & Waite, J. H. (2009) Collagen insulated from tensile damage by domains that unfold reversibly: In situ X-ray investigation of mechanical yield and damage repair in the mussel byssus. *J. Struct. Biol.* **167**, 47-54.
- [112] Tanaka, K., Inoue, S. & Mizuno, S. (1999) Hydrophobic interaction of P25, containing Asn-linked oligosaccharide chains, with the H-L complex of silk fibroin produced by *Bombyx mori*. *Insect Biochem. Mol. Biol.* **29**, 269-276.
- [113] Hagn, F., Thamm, C., Scheibel, T. & Kessler, H. (2011) pH-dependent dimerization and salt-dependent stabilization of the N-terminal domain of spider dragline silk - implications for fiber formation. *Angew. Chem. Int. Ed. Engl.* **50**, 310-313.
- [114] Gaines, W. A., Sehorn, M. G. & Marcotte, W. R., Jr. (2010) Spidroin N-terminal domain promotes a pH-dependent association of silk proteins during self-assembly. *J. Biol. Chem.* **285**, 40745-40753.
- [115] Dicko, C., Kenney, J. M., Knight, D. & Vollrath, F. (2004) Transition to a  $\beta$ -sheet-rich structure in spidroin in vitro: The effects of pH and cations. *Biochemistry-US* **43**, 14080-14087.
- [116] Eisoltd, L., Thamm, C. & Scheibel, T. (2012) The role of terminal domains during storage and assembly of spider silk proteins. *Biopolymers* **97**, 355-361.
- [117] Landreh, M. *et al.* (2010) A pH-dependent dimer lock in spider silk protein. *J. Mol. Biol.* **404**, 328-336.
- [118] Hagn, F. *et al.* (2010) A conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature* **465**, 239-243.
- [119] Eisoltd, L., Hardy, J. G., Heim, M. & Scheibel, T. R. (2010) The role of salt and shear on the storage and assembly of spider silk proteins. *J. Struct. Biol.*, 413-419.
- [120] Exler, J. H., Hummerich, D. & Scheibel, T. (2007) The amphiphilic properties of spider silks are important for spinning. *Angew. Chem. Int. Ed. Engl.* **46**, 3559-3562.
- [121] Harrington, M. J., Masic, A., Holten-Andersen, N., Waite, J. H. & Fratzl, P. (2010) Iron-clad fibers: a metal-based biological strategy for hard flexible coatings. *Science* **328**, 216-220.
- [122] Brown, C. H. (1952) Some structural proteins of *Mytilus edulis*. *Q. J. Microsc. Sci.* **93**, 487-503.
- [123] Cook, M. Composition of mussel and barnacle deposits at the attachment interface in *Adhesion in biological systems* (Editor Manly, R. S.), Academic Press (1970), 139-150.
- [124] Bella, J., Liu, J., Kramer, R., Brodsky, B. & Berman, H. M. (2006) Conformational effects of Gly-X-Gly interruptions in the collagen triple helix. *J. Mol. Biol.* **362**, 298-311.
- [125] Hagenau, A. (2011) *Analyse der Struktur-Funktionsbeziehungen natürlicher Muschelbyssusfäden der Miesmuschel Mytilus galloprovincialis*, Dissertation, Department Chemie, TU München.
- [126] Krauss, S., Metzger, T. H., Fratzl, P. & Harrington, M. J. (2013) Self-repair of a biological fiber guided by an ordered elastic framework. *Biomacromolecules* **14**, 1520-1528.
- [127] Vrhovski, B. & Weiss, A. S. (1998) Biochemistry of tropoelastin. *Eur. J. Biochem.* **258**, 1-18.

- 
- [128] Rauscher, S. & Pomes, R. (2012) Structural disorder and protein elasticity. *Adv. Exp. Med. Biol.* **725**, 159-183.
  - [129] Rosenbloom, J., Abrams, W. R. & Mecham, R. (1993) Extracellular matrix 4: the elastic fiber. *Faseb J.* **7**, 1208-1218.
  - [130] Ringli, C., Keller, B. & Ryser, U. (2001) Glycine-rich proteins as structural components of plant cell walls. *Cell. Mol. Life Sci.* **58**, 1430-1441.
  - [131] Condit, C. M. & Meagher, R. B. (1986) A Gene Encoding a Novel Glycine-Rich Structural Protein of Petunia. *Nature* **323**, 178-181.
  - [132] Condit, C. M. & Meagher, R. B. (1987) Expression of a Gene Encoding a Glycine-Rich Protein in Petunia. *Mol. Cell. Biol.* **7**, 4273-4279.
  - [133] Carrington, E. & Gosline, J. M. (2004) Mechanical design of mussel byssus: Load cycle and strain rate dependence. *Am. Malacol. Bull.* **18**, 135-142.
  - [134] Guinea, G. V., Elices, M., Real, J. I., Gutierrez, S. & Perez-Rigueiro, J. (2005) Reproducibility of the tensile properties of spider (*Argiope trifasciata*) silk obtained by forced silking. *J. Exp. Zool. Part A* **303A**, 37-44.
  - [135] Work, R. W. & Emerson, P. D. (1982) An Apparatus and Technique for the Forcible Silking of Spiders. *J. Arachnol.* **10**, 1-10.
  - [136] Marsano, E. *et al.* (2005) Wet spinning of *Bombyx mori* silk fibroin dissolved in *N*-methyl morpholine *N*-oxide and properties of regenerated fibres. *Int. J. Biol. Macromol.* **37**, 179-188.
  - [137] Matsumoto, K. & Uejima, H. (1997) Regenerated protein fibers. 1. Research and development of a novel solvent for silk fibroin. *J. Polym. Sci. Pol. Chem.* **35**, 1949-1954.
  - [138] Rodbumrer, P. *et al.* (2012) Functional expression of a *Bombyx mori* cocoonase: potential application for silk degumming. *Acta Biochim. Biophys. Sin. (Shanghai)* **44**, 974-983.
  - [139] Heidebrecht, A. & Scheibel, T. (2013) Recombinant production of spider silk proteins. *Adv. Appl. Microbiol.* **82**, 115-153.
  - [140] Asakura, T., Tanaka, C., Yang, M., Yao, J. & Kurokawa, M. (2004) Production and characterization of a silk-like hybrid protein, based on the polyalanine region of *Samia cynthia ricini* silk fibroin and a cell adhesive region derived from fibronectin. *Biomaterials* **25**, 617-624.
  - [141] Teule, F. *et al.* (2012) Silkworms transformed with chimeric silkworm/spider silk genes spin composite silk fibers with improved mechanical properties. *Proc. Natl. Acad. Sci. USA* **109**, 923-928.
  - [142] Yang, M. *et al.* (2008) Silklike materials constructed from sequences of *Bombyx mori* silk fibroin, fibronectin, and elastin. *J. Biomed. Mater. Res. A* **84**, 353-363.
  - [143] Pakkanen, O., Pirskanen, A. & Myllyharju, J. (2006) Selective expression of nonsecreted triple-helical and secreted single-chain recombinant collagen fragments in the yeast *Pichia pastoris*. *J. Biotechnol.* **123**, 248-256.
  - [144] Hellio, C., Bourgougnon, N. & Le Gal, Y. (2000) Phenoloxidase (EC 1.14.18.1) from the byssus gland of *Mytilus edulis*: Purification, partial characterization and application for screening products with potential antifouling activities. *Biofouling* **16**, 235-244.
  - [145] Waite, J. H. (1985) Catechol Oxidase in the Byssus of the Common Mussel, *Mytilus Edulis* L. *J. Mar. Biol. Assoc. Uk* **65**, 359-371.
  - [146] Vitellarro-Zuccarello, L. (1981) Ultrastructural and Cytochemical Study on the Enzyme Gland of the Foot of a Mollusk. *Tissue Cell* **13**, 701-713.
  - [147] Elsner, M. B. (2011) *Herstellung, Reinigung und Analyse flankierender Regionen eines M. galloprovincialis Präkollagen-Moleküls in E. coli*, Masterarbeit, Fakultät für Biologie, Chemie und Geowissenschaften, Universität Bayreuth.
  - [148] Huemmerich, D. *et al.* (2004) Primary structure elements of spider dragline silks and their contribution to protein solubility. *Biochemistry-Us* **43**, 13604-13612.

- [149] Heim, M., Ackerschott, C. B. & Scheibel, T. (2010) Characterization of recombinantly produced spider flagelliform silk domains. *J. Struct. Biol.* **170**, 420-425.
- [150] Niedermeyer, T. H. & Strohal, M. (2012) mMass as a software tool for the annotation of cyclic peptide tandem mass spectra. *PLoS One* **7**, e44913.
- [151] Snider, C., Jayasinghe, S., Hristova, K. & White, S. H. (2009) MPEx: a tool for exploring membrane proteins. *Protein Sci.* **18**, 2624-2628.
- [152] Kyte, J. & Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132.
- [153] Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res.* **22**, 4673-4680.
- [154] Smith, T. F. & Waterman, M. S. (1981) Identification of common molecular subsequences. *J. Mol. Biol.* **147**, 195-197.
- [155] Huang, X. Q. & Miller, W. (1991) A Time-Efficient, Linear-Space Local Similarity Algorithm. *Adv. Appl. Math.* **12**, 337-357.
- [156] Novoa, E. M. & de Pouplana, L. R. (2012) Speeding with control: codon usage, tRNAs, and ribosomes. *Trends Genet.* **28**, 574-581.
- [157] Hershberg, R. & Petrov, D. A. (2008) Selection on Codon Bias. *Annu. Rev. Genet.* **42**, 287-299.
- [158] Fuhrmann, M. *et al.* (2004) Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. *Plant Mol. Biol.* **55**, 869-881.
- [159] Sharp, P. M. & Li, W. H. (1987) The codon Adaptation Index - a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **15**, 1281-1295.
- [160] Stellwagen, N. C. (2009) Electrophoresis of DNA in agarose gels, polyacrylamide gels and in free solution. *Electrophoresis* **30**, S188-S195.
- [161] Görg, A. & Westermeier, R. Elektrophoretische Verfahren in *Bioanalytik* (Editoren Lottspeich, F. & Zorbas, H.), Spektrum Akademischer Verlag GmbH (1998), 217-252.
- [162] Mullis, K. B. & Faloona, F. A. (1987) Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Meth. Enzymol.* **155**, 335-350.
- [163] Sambrook, J. & Russell, D. W. Molecular cloning: a laboratory manual. 3. Edition, Cold Spring Harbor Laboratory Press (2001).
- [164] Porath, J., Carlsson, J., Olsson, I. & Belfrage, G. (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598-599.
- [165] Sulkowski, E. (1985) Purification of Proteins by IMAC. *Trends Biotechnol.* **3**, 1-7.
- [166] Evans, T. C., Jr., Benner, J. & Xu, M. Q. (1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* **7**, 2256-2264.
- [167] Southworth, M. W., Amaya, K., Evans, T. C., Xu, M. Q. & Perler, F. B. (1999) Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* gyrase A intein. *Biotechniques* **27**, 110-120.
- [168] Mossessova, E. & Lima, C. D. (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol. Cell* **5**, 865-876.
- [169] Butt, T. R., Edavettal, S. C., Hall, J. P. & Mattern, M. R. (2005) SUMO fusion technology for difficult-to-express proteins. *Protein Expres. Purif.* **43**, 1-9.
- [170] Reynolds, J. A. & Tanford, C. (1970) Binding of dodecyl sulfate to proteins at high binding ratios. Possible implications for the state of proteins in biological membranes. *Proc. Natl. Acad. Sci. USA* **66**, 1002-1007.

- [171] Davis, B. J. (1964) Disc Electrophoresis - II. Method and Application to Human Serum Proteins. *Ann. N. Y. Acad. Sci.* **121**, 404-427.
- [172] Ornstein, L. (1964) Disc Electrophoresis - I. Background and Theory. *Ann. N. Y. Acad. Sci.* **121**, 321-349.
- [173] Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- [174] Schägger, H. (2006) Tricine-SDS-PAGE. *Nat. Protoc.* **1**, 16-22.
- [175] Schägger, H. & von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368-379.
- [176] Wray, W., Boulikas, T., Wray, V. P. & Hancock, R. (1981) Silver staining of proteins in polyacrylamide gels. *Anal. Biochem.* **118**, 197-203.
- [177] Switzer, R. C., 3rd, Merrill, C. R. & Shifrin, S. (1979) A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. *Anal. Biochem.* **98**, 231-237.
- [178] Mäntele, W. Spektroskopie in *Bioanalytik* (Editoren Lottspeich, F. & Zorbas, H.), Spektrum Akademischer Verlag GmbH (1998), 131-178.
- [179] Gallagher, S. R. Quantitation of DNA and RNA with absorption and fluorescence spectroscopy in *Current Protocols in Molecular Biology* (Editoren Ausubel, F. A. et al.), John Wiley & Sons (1989), A.3.9-A3.15.
- [180] Glasel, J. A. (1995) Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. *Biotechniques* **18**, 62-63.
- [181] Gill, S. C. & von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.
- [182] Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411-2423.
- [183] Bulheller, B. M., Rodger, A. & Hirst, J. D. (2007) Circular and linear dichroism of proteins. *Phys. Chem. Chem. Phys.* **9**, 2020-2035.
- [184] Welfle, H. Circular dichroismus in *Methoden in der Proteinanalytik* (Editor Holtzhauer, M.), Springer, Berlin (1996), 103-123.
- [185] Corrêa, D. H. A. & Ramos, C. H. I. (2009) The use of circular dichroism spectroscopy to study protein folding, form and function. *African Journal of Biochemistry Research* **3**, 164-173.
- [186] Bujacz, A. (2012) Structures of bovine, equine and leporine serum albumin. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 1278-1289.
- [187] Mäntele, W. Fluoreszenzspektroskopie in *Bioanalytik* (Editoren Lottspeich, F. & Zorbas, H.), Springer Akademischer Verlag GmbH (1998), 167-172.
- [188] Schmid, F. X. Spectral methods of characterizing protein conformation and conformational changes in *Protein Structure: A Practical Approach* (Editor Creighton, T. E.), Oxford University Press (1997).
- [189] Ladokhin, A. S. Fluorescence Spectroscopy in Peptide and Protein Analysis in *Encyclopedia of Analytical Chemistry* (Editor Meyers, R. A.), John Wiley & Sons Ltd. (2000), 5762-5779.
- [190] Yates, J. R., III (1998) Mass spectrometry and the age of the proteome. *J. Mass. Spectrom.* **33**, 1-19.
- [191] Rivers, R. C. et al. (2008) Molecular determinants of the aggregation behavior of alpha- and beta-synuclein. *Protein Sci.* **17**, 887-898.
- [192] Dominguez, A., Fernandez, A., Gonzalez, N., Iglesias, E. & Montenegro, L. (1997) Determination of critical micelle concentration of some surfactants by three techniques. *J. Chem. Educ.* **74**, 1227-1231.

- [193] White, S. H. & Wimley, W. C. (1999) Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 319-365.
- [194] Mitchell, D. J. & Ninham, B. W. (1981) Micelles, Vesicles and Micro-Emulsions. *J. Chem. Soc. Faraday Trans. 2* **77**, 601-629.
- [195] Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**, 195-202.
- [196] Marblestone, J. G. *et al.* (2006) Comparison of SUMO fusion technology with traditional gene fusion systems: Enhanced expression and solubility with SUMO. *Protein Sci.* **15**, 182-189.
- [197] Saitoh, H., Uwada, J. & Azusa, K. (2009) Strategies for the expression of SUMO-modified target proteins in *Escherichia coli*. *Methods Mol. Biol.* **497**, 211-221.
- [198] Rippon, W. B. & Walton, A. G. (1971) Optical properties of the polyglycine II helix. *Biopolymers* **10**, 1207-1212.
- [199] Tompa, P. (2012) Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* **37**, 509-516.
- [200] Uversky, V. N. (2011) Intrinsically disordered proteins from A to Z. *Int. J. Biochem. Cell Biol.* **43**, 1090-1103.
- [201] Mao, A. H., Lyle, N. & Pappu, R. V. (2013) Describing sequence-ensemble relationships for intrinsically disordered proteins. *Biochem. J.* **449**, 307-318.
- [202] Babu, M. M., Kriwacki, R. W. & Pappu, R. V. (2012) Structural biology. Versatility from protein disorder. *Science* **337**, 1460-1461.
- [203] Das, R. K., Mittal, A. & Pappu, R. V. (2013) How is functional specificity achieved through disordered regions of proteins? *Bioessays* **35**, 17-22.
- [204] Pujol, J. P. (1967) Formation of the Byssus in the Common Mussel (*Mytilus edulis* L.). *Nature* **214**, 204-205.
- [205] Pujol, J. P., Houvenaghel, G. & Bouillon, J. (1972) Le collagène du byssus de *Mytilus edulis* L. I. Ultrastructure des cellules sécrétrices. *Archives de Zoologie Expérimentale et Generale* **113**, 251-264.
- [206] Wimley, W. C. *et al.* (1998) Folding of beta-sheet membrane proteins: a hydrophobic hexapeptide model. *J. Mol. Biol.* **277**, 1091-1110.
- [207] Wimley, W. C. & White, S. H. (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol.* **3**, 842-848.
- [208] Swartz, J. R. (2001) Advances in *Escherichia coli* production of therapeutic proteins. *Curr. Opin. Biotech.* **12**, 195-201.
- [209] Baneyx, F. (1999) Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotech.* **10**, 411-421.
- [210] Zerbs, S., Frank, A. M. & Collart, F. R. (2009) Bacterial Systems for Production of Heterologous Proteins. *Method. Enzymol.* **463**, 149-168.
- [211] Dyson, M. R., Shadbolt, S. P., Vincent, K. J., Perera, R. L. & McCafferty, J. (2004) Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression. *BMC Biotechnol.* **4**, DOI: 10.1186/1472-6750-4-32
- [212] Mergulhao, F. J. M., Monteiro, G. A., Cabral, J. M. S. & Taipa, M. A. (2004) Design of bacterial vector systems for the production of recombinant proteins in *Escherichia coli*. *J. Microbiol. Biotechnol.* **14**, 1-14.
- [213] Graslund, S. *et al.* (2008) Protein production and purification. *Nat. Methods* **5**, 135-146.
- [214] Bauer, F. & Scheibel, T. (2012) Artificial Egg Stalks Made of a Recombinantly Produced Lacewing Silk Protein. *Angew. Chem. Int. Ed.* **51**, 6521-6524.



- [215] Fahnestock, S. R. & Irwin, S. L. (1997) Synthetic spider dragline silk proteins and their production in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **47**, 23-32.
- [216] Fahnestock, S. R., Yao, Z. & Bedzyk, L. A. (2000) Microbial production of spider silk proteins. *J. Biotechnol.* **74**, 105-119.
- [217] Fukushima, Y. (1998) Genetically engineered syntheses of tandem repetitive polypeptides consisting of glycine-rich sequence of spider dragline silk. *Biopolymers* **45**, 269-279.
- [218] Prince, J. T., Mcgrath, K. P., Digirolamo, C. M. & Kaplan, D. L. (1995) Construction, Cloning, and Expression of Synthetic Genes Encoding Spider Dragline Silk. *Biochemistry-Us* **34**, 10879-10885.
- [219] Zhou, Y. T., Wu, S. X. & Conticello, V. P. (2001) Genetically directed synthesis and spectroscopic analysis of a protein polymer derived from a flagelliform silk sequence. *Biomacromolecules* **2**, 111-125.
- [220] Plotkin, J. B. & Kudla, G. (2011) Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet.* **12**, 32-42.
- [221] Gingold, H. & Pilpel, Y. (2011) Determinants of translation efficiency and accuracy. *Mol. Syst. Biol.* **7**, DOI: 10.1038/msb.2011.14
- [222] Grantham, R., Gautier, C., Gouy, M., Jacobzone, M. & Mercier, R. (1981) Codon catalog usage is a genome strategy modulated for gene expressivity. *Nucleic Acids Res.* **9**, R43-R74.
- [223] Bennetzen, J. L. & Hall, B. D. (1982) Codon Selection in Yeast. *J. Biol. Chem.* **257**, 3026-3031.
- [224] Gouy, M. & Gautier, C. (1982) Codon Usage in Bacteria - Correlation with Gene Expressivity. *Nucleic Acids Res.* **10**, 7055-7074.
- [225] Wootton, J. C. (1994) Sequences with 'unusual' amino acid compositions. *Curr. Opin. Struct. Biol.* **4**, 413-421.
- [226] Chong, S., Williams, K. S., Wotkowicz, C. & Xu, M. Q. (1998) Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar membrane ATPase intein. *J. Biol. Chem.* **273**, 10567-10577.
- [227] Chong, S. *et al.* (1997) Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **192**, 271-281.
- [228] New England Biolabs (2008) IMPACT™-Kit: Protein Purification System Featuring Fusion to C- or N-Terminus of the Target Protein. *NEW ENGLAND BioLabs Inc. Instruction Manual Version 1.1*
- [229] Kelly, S. M., Jess, T. J. & Price, N. C. (2005) How to study proteins by circular dichroism. *Biochim. Biophys. Acta* **1751**, 119-139.
- [230] Bykov, S. & Asher, S. (2010) Raman Studies of Solution Polyglycine Conformations. *J. Phys. Chem.* **114**, 6636-6641.
- [231] Cubellis, M. V., Caillez, F., Blundell, T. L. & Lovell, S. C. (2005) Properties of polyproline II, a secondary structure element implicated in protein-protein interactions. *Proteins* **58**, 880-892.
- [232] Teufel, D. P., Johnson, C. M., Lum, J. K. & Neuweiler, H. (2011) Backbone-driven collapse in unfolded protein chains. *J. Mol. Biol.* **409**, 250-262.
- [233] Tran, H. T., Mao, A. & Pappu, R. V. (2008) Role of backbone-solvent interactions in determining conformational equilibria of intrinsically disordered proteins. *J. Am. Chem. Soc.* **130**, 7380-7392.
- [234] Graham, L. A. & Davies, P. L. (2005) Glycine-rich antifreeze proteins from snow fleas. *Science* **310**, 461.
- [235] Pentelute, B. L. *et al.* (2008) X-ray structure of snow flea antifreeze protein determined by racemic crystallization of synthetic protein enantiomers. *J. Am. Chem. Soc.* **130**, 9695-9701.
- [236] Seelig, J. (2004) Thermodynamics of lipid-peptide interactions. *Biochim. Biophys. Acta* **1666**, 40-50.
- [237] Terzi, E., Holzemann, G. & Seelig, J. (1995) Self-association of beta-amyloid peptide (1-40) in solution and binding to lipid membranes. *J. Mol. Biol.* **252**, 633-642.

- [238] Welker, S. *et al.* (2010) Hsp12 is an intrinsically unstructured stress protein that folds upon membrane association and modulates membrane function. *Mol. Cell* **39**, 507-520.
- [239] Wieprecht, T., Apostolov, O., Beyermann, M. & Seelig, J. (1999) Thermodynamics of the alpha-helix-coil transition of amphipathic peptides in a membrane environment: implications for the peptide-membrane binding equilibrium. *J. Mol. Biol.* **294**, 785-794.
- [240] Wieprecht, T. *et al.* (1996) Conformational and functional study of magainin 2 in model membrane environments using the new approach of systematic double-D-amino acid replacement. *Biochemistry-Us* **35**, 10844-10853.
- [241] Heim, M., Elsner, M. B. & Scheibel, T. (2013) Lipid-Specific beta-Sheet Formation in a Mussel Byssus Protein Domain. *Biomacromolecules* **14**, 3238-3245.
- [242] Ladokhin, A. S., Fernandez-Vidal, M. & White, S. H. (2010) CD spectroscopy of peptides and proteins bound to large unilamellar vesicles. *J. Membr. Biol.* **236**, 247-253.
- [243] White, S. H., Ladokhin, A. S., Jayasinghe, S. & Hristova, K. (2001) How membranes shape protein structure. *J. Biol. Chem.* **276**, 32395-32398.
- [244] Israelachvili, J. N. & Mitchell, D. J. (1975) A model for the packing of lipids in bilayer membranes. *Biochim. Biophys. Acta* **389**, 13-19.
- [245] King, M. D. & Marsh, D. (1987) Head Group and Chain-Length Dependence of Phospholipid Self-Assembly Studied by Spin-Label Electron-Spin-Resonance. *Biochemistry-Us* **26**, 1224-1231.
- [246] Moreira, L. & Firoozabadi, A. (2010) Molecular thermodynamic modeling of specific ion effects on micellization of ionic surfactants. *Langmuir* **26**, 15177-15191.
- [247] Jacobson, K. & Papahadjopoulos, D. (1975) Phase transitions and phase separations in phospholipid membranes induced by changes in temperature, pH, and concentration of bivalent cations. *Biochemistry-Us* **14**, 152-161.
- [248] Spinozzi, F., Paccamiccio, L., Mariani, P. & Amaral, L. Q. (2010) Melting regime of the anionic phospholipid DMPG: new lamellar phase and porous bilayer model. *Langmuir* **26**, 6484-6493.
- [249] Hermanson, K. D., Huemmerich, D., Scheibel, T. & Bausch, A. R. (2007) Engineered microcapsules fabricated from reconstituted spider silk. *Adv. Mater.* **19**, 1810-1815.
- [250] Schulz, G. E. (2000) beta-Barrel membrane proteins. *Curr. Opin. Struct. Biol.* **10**, 443-447.
- [251] Wimley, W. C. (2002) Toward genomic identification of beta-barrel membrane proteins: composition and architecture of known structures. *Protein Sci.* **11**, 301-312.
- [252] Jacobs, R. E. & White, S. H. (1989) The nature of the hydrophobic binding of small peptides at the bilayer interface: implications for the insertion of transbilayer helices. *Biochemistry-Us* **28**, 3421-3437.
- [253] Tamarin, A. & Keller, P. J. (1972) An ultrastructural study of the byssal thread forming system in *Mytilus*. *J. Ultrastruct. Res.* **40**, 401-416.
- [254] Vitellarro-Zuccarello, L. (1980) The collagen gland of *Mytilus galloprovincialis*: an ultrastructural and cytochemical study on secretory granules. *J. Ultrastruct. Res.* **73**, 135-147.
- [255] Fraser, R. D., MacRae, T. P. & Suzuki, E. (1979) Chain conformation in the collagen molecule. *J. Mol. Biol.* **129**, 463-481.
- [256] Rich, A. & Crick, F. H. (1961) The molecular structure of collagen. *J. Mol. Biol.* **3**, 483-506.
- [257] Vaccaro, E. & Waite, J. H. (2001) Yield and post-yield behavior of mussel byssal thread: A self-healing biomolecular material. *Biomacromolecules* **2**, 906-911.
- [258] Harrington, M. J. *et al.* (2012) Pseudoelastic behaviour of a natural material is achieved via reversible changes in protein backbone conformation. *J. R. Soc. Interface* **9**, 2911-2922.
- [259] Miserez, A., Wasko, S. S., Carpenter, C. F. & Waite, J. H. (2009) Non-entropic and reversible long-range deformation of an encapsulating bioelastomer. *Nat. Mater.* **8**, 910-916.

- [260] Fullenkamp, D. E., He, L. H., Barrett, D. G., Burghardt, W. R. & Messersmith, P. B. (2013) Mussel-Inspired Histidine-Based Transient Network Metal Coordination Hydrogels. *Macromolecules* **46**, 1167-1174.
- [261] Krogsgaard, M., Behrens, M. A., Pedersen, J. S. & Birkedal, H. (2013) Self-Healing Mussel-Inspired Multi-pH-Responsive Hydrogels. *Biomacromolecules* **14**, 297-301.
- [262] Vatankhah-Varnoosfaderani, M., GhavamiNejad, A., Hashmi, S. & Stadler, F. J. (2013) Mussel-inspired pH-triggered reversible foamed multi-responsive gel - the surprising effect of water. *Chem Commun. (Camb.)* **49**, 4685-4687.



## 8 Abkürzungen

Dieses Verzeichnis bezieht sich auf häufig in Text, Abbildungen und Tabellen verwendete Abkürzungen. Zur Abkürzung von Chemikalien siehe Kapitel 2.1.1!

A, Ala (Aminosäure)	Alanin
$[\theta]_{\text{MRW}}$	mittlere molare Elliptizität pro Aminosäure
A (Nukleosid)	Adenosin
<i>A. diadematus</i>	<i>Araneus diadematus</i> (Gartenkreuzspinne)
$A_{254}$	Wert der Absorption von UV-Licht der Wellenlänge 254 nm
$A_{280}$	Wert der Absorption von UV-Licht der Wellenlänge 280 nm
ADF	<i>Araneus diadematus</i> Fibroin
AMW	artifizielles Meerwasser
<i>B. mori</i>	<i>Bombyx mori</i> (Maulbeerspinner)
BLAST	<i>Basic Local Alignment Search Tool</i>
bp	Basenpaar(e)
BSA	bovines Serumalbumin
c	Konzentration
C (Nukleosid)	Cytidin
C, Cys (Aminosäure)	Cystein
CBD	Chitinbindedomäne
CD	Zirkulardichroismus ( <i>circular dichroism</i> )
cDNA	komplementäre ( <i>complementary</i> ) DNA
CIP	<i>cleaning-in-place</i>
cwCT	pflanzenzellwandähnliche, carboxyterminal flankierende Region
cwNT	pflanzenzellwandähnliche, aminoterminal flankierende Region
d	Schichtdicke
D, Asp	Asparaginsäure
Da	atomare Masseneinheit ( $\sim 1,66 \cdot 10^{-24}$ g), $1 \text{ Da} \triangleq 1 \text{ g/mol}$
DNA	Desoxyribonukleinsäure
dNTP	2'-Desoxyribonukleosid-5'-triphosphat
DOPA	L-3,4-Dihydroxyphenylalanin
E, Glu	Glutaminsäure
<i>E. coli</i>	<i>Escherichia coli</i> (gramnegatives Bakterium)
<i>E. australis</i>	<i>Euprostenops australis</i>
eADF	<i>engineered Araneus diadematus</i> Fibroin
eCT	elastinähnliche, carboxyterminal flankierende Region

## 8 Abkürzungen

---

eNT	elastinähnliche, aminoterminal flankierende Region
F, Phe	Phenylalanin
FITC	Fluorescein-Isothiocyanat
FPLC	<i>fast protein liquid chromatography</i>
fwd	Kennzeichnung für Vorwärtsprimer ( <i>forward primer</i> )
g	Schwerebeschleunigung der Erde (9,81 m/s <sup>2</sup> )
G (Nukleosid)	Guanosin
G, Gly (Aminosäure)	Glycin
GPC	Gelpermeationschromatographie
GRP	glycinreiche Proteine (aus Pflanzenzellwänden)
H, His	Histidin
I, Ile	Isoleucin
IDA	Iminodiessigsäure
IDP	intrinsisch entfaltetes Protein ( <i>intrinsically disordered protein</i> )
IMAC	<i>immobilized metal affinity chromatography</i>
K, Lys	Lysin
kb	Kilobasen
L, Leu	Leucin
<i>L. hesperus</i>	<i>Latrodectus hesperus</i> (Westliche Schwarze Witwe)
LB	komplexes Nährmedium zur Bakterienkultivierung ( <i>lysogeny broth</i> )
M (als Konzentration)	1 M = 1 mol/l
M, Met (Aminosäure)	Methionin
MAS	Seide der Großen Ampullendrüse ( <i>major ampullate silk</i> )
MaSp	Spidroin(e) der Großen Ampullendrüse ( <i>major ampullate spidroin</i> )
Mfp	Muschelfußprotein ( <i>mussel foot protein</i> )
MiSp	Spidroin(e) der Kleinen Ampullendrüse ( <i>minor ampullate spidroin</i> )
MQ	Milli-Q Wasser (Reinstwasser)
M <sub>w</sub>	Molekulargewicht
Mxe, <i>M. xenopi</i>	<i>Mycobacterium xenopi</i>
n	Anzahl der Peptidbindungen
N, Asn	Asparagin
<i>N. clavipes</i>	<i>Nephila clavipes</i> (Goldene Seidenspinne)
n.b.	nicht bestimmbar / nicht bestimmt
N <sub>A</sub>	Avogadrokonstante ( $\sim 6,022141 \cdot 10^{23} \text{ mol}^{-1}$ )
NMR	Kernspinresonanz ( <i>nuclear magnetic resonance</i> )
NRC	nichtrepetitiver C-Terminus
NRN	nichtrepetitiver N-Terminus
OD <sub>600</sub>	optische Dichte bei 600 nm Wellenlänge
P, Pro	Prolin
p. a.	<i>pro analysis</i>

---

PCR	Polymerasekettenreaktion ( <i>polymerase chain reaction</i> )
PFC	<i>partitioning-folding coupling</i>
PGII	Polyglycin-II
PPII	Polyprolin-II
PSK	Polystyrolkugeln
Q, Gln	Glutamin
r	Radius
R, Arg	Arginin
rev	Kennzeichnung für Rückwärtsprimer ( <i>reverse primer</i> )
RT	Raumtemperatur
S, Ser	Serin
sCT	seidenähnliche, carboxyterminal flankierende Region
Sfl	synthetisches Flagelliformprotein
sNT	seidenähnliche, aminoterminal flankierende Region
SOC	komplexes Nährmedium zur Bakterienkultivierung
SOP	Standardarbeitsanweisung ( <i>standard operating procedure</i> )
SUMO	<i>small ubiquitin-like modifier</i>
SUV	kleine(s) unilamellare(s) (Lipid)Vesikel ( <i>small unilamellar vesicle(s)</i> )
SV	Säulenvolumen
T (Aminosäure)	Threonin
T (Nukleosid)	Tymin
TAE	Tris-Acetat-EDTA-Puffer
TMP	Fadenmatrixprotein ( <i>thread matrix protein</i> )
U	Umdrehungen
u	<i>unit</i> (Maß der Enzymaktivität, nicht einheitlich definiert)
UV	ultraviolett
V	Volumen
V, Val (Aminosäure)	Valin
v/v	Volumenprozent ( <i>volume per volume</i> )
VE	vollentsalzt
VIS	visuell
W, Trp	Tryptophan
w/v	Gewichtsprozent ( <i>weight per volume</i> )
Y, Tyr	Tyrosin
$\epsilon_{280\text{ nm}}$	Molarer Extinktionskoeffizient [ $\text{l}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ ] bei 280 nm Wellenlänge
$\theta$	gemessene Elliptizität
$\rho$	Dichte





## 9 Publikationen

Heim, M., Elsner, M. B. & Scheibel, T. (2013) Lipid-Specific beta-Sheet Formation in a Mussel Byssus Protein Domain. *Biomacromolecules* **14**, 3238-3245.



## 10 Danksagung

Die hier vorliegende Arbeit wurde von März 2009 bis Februar 2014 am Lehrstuhl Biomaterialien der Universität Bayreuth angefertigt.

Mein besonderer Dank gilt meinem Doktorvater Prof. Dr. Thomas Scheibel, der es mir ermöglichte, dieses interessante, herausfordernde und spannende Thema an seinem Lehrstuhl bearbeiten zu können. Darüber hinaus möchte ich ihm auch dafür danken, dass er als wissenschaftlicher Ansprechpartner stets ein offenes Ohr hatte und durch zahlreiche Diskussionen dazu beitrug, die kleinen und großen Herausforderungen, die das Thema mit sich brachte, mit einer über den Tellerrand hinausgehenden Sichtweise erfolgreich meistern zu können. Darüber hinaus möchte ich mich bei Prof. Dr. Thomas Scheibel, der Deutschen Forschungsgemeinschaft und der Graduiertenförderung der Universitäten Bayern e.V. auch für die finanzielle Unterstützung bedanken, die eine erfolgreiche Bearbeitung überhaupt erst ermöglichte.

Ferner möchte ich mich natürlich auch herzlich bei Herrn Prof. Dr. Franz X. Schmid und Herrn Dr. Stefan Geimer dafür bedanken, dass sie sich bereit erklärten, den Fortgang der hier vorliegenden Arbeit als engagierte Mentoren des Graduiertenprogramms „Molekulare Biowissenschaften“ zu begleiten, und dass sie mir stets mit Rat und Tat zur Seite standen und mir so auch neue Blickwinkel eröffnen konnten.

Ein herzliches Dankeschön geht natürlich an all die Helfer, die mir bei der Durchführung von Versuchen und der Auswertung und Interpretation von Daten hilfreich zur Seite standen:

- an Dr. Michael H. Suhre und Adrian Golser für die Bereitstellung von aus der cDNA-Bank isolierten preCol-Klonen
- an Dr. Lukas Eisoldt für die Bereitstellung eines SUMO-Expressionsvektors sowie der Ulp1-Protease
- an Dr. Claudia Blüm und Dr. Christian Ackerschott für die Überlassung der als Kontrolle verwendeten Spinnenseidenproteine eADF4(C16) und Sfl
- an Johannes Diehl und Andreas Schmidt für ihre Hilfe bei den Fermentationen

- an meine frühere Masterstudentin und jetzige Kollegin Martina B. Elsner für ihre Mitarbeit bei der Entwicklung der Strategie zur rekombinanten Produktion und Reinigung der *flank*-Proteine
- an Dr. Martin Humenik und Christopher Synatschke für die Vorbereitung und Durchführung der MALDI-TOF Analysen
- an Jasmin Wickinghoff für ihre Unterstützung bei Konierungsarbeiten

Bedanken möchte ich mich zudem bei allen Mitgliedern meiner Subgroup „Aquatische Fasern“ für zahlreiche und konstruktive Diskussionen und Vorschläge.

Ein herzliches Dankeschön gilt ferner meinen Bürokollegen Dr. Anja Hagenau, Dr. Kristina Spieß, Dr. Felix Bauer, Dr. Claudia Blüm, Aniola Heidebrecht, Elena Dobelhofer und Dr. Ute Slotta für eine konstruktive und freundschaftliche Arbeitsatmosphäre sowie für ihre stete Bereitschaft, wissenschaftliche Fragestellungen umfassend zu diskutieren.

Für die nette Aufnahme nach meinem Umzug in die Ludwig-Thoma-Straße möchte ich mich zudem bei Aniola Heidebrecht, Gregor Lang, Johannes Diehl und Stephan Jokisch bedanken.

Ferner bedanke ich mich bei unserer Sekretärin Susanne Schramm dafür, dass sie einem bei organisatorischen Dingen immer freundlich zur Seite stand und einen souverän durch die Verwaltungsebenen des öffentlichen Dienstes manövrierte, sowie bei allen technischen Assistenten für die gute Organisation und den reibungslosen Ablauf des Laboralltags.

Ein herzlicher Dank auch noch an die (Hobby-)Fußballer des Lehrstuhls, an die Mitglieder der Schafkopfrunde sowie an alle Mitarbeiter des Lehrstuhls für die zahlreichen extrakurrikularen Aktivitäten, die dem „Allgäuer Exil-Franken“ das Leben in der „Neuen Heimat“ sehr angenehm gestalteten und nicht zuletzt dazu beitrugen, dass er sich immer gerne an diese Zeit zurückerinnern wird.

Abschließend gebührt mein allergrößter Dank meiner Frau Carmen und meiner Familie, die mich immerwährend unterstützt, an mich geglaubt und mir den Rücken freigehalten haben, so dass ich mich meiner Promotion stets mit vollem Einsatz widmen konnte.

## 11 Anhang

### 11.1 Verwendete Proteine

Im Folgenden sind Aminosäuresequenzen, molare Extinktionskoeffizienten, Molekulargewichte und (theoretische) isoelektrische Punkte (letztere drei alle mit ProtParam berechnet, vgl. Kapitel 2.2.1) der im Rahmen dieser Arbeit erzeugten *flank*-Proteine, des SUMO- und CBD-tags sowie der verwendeten Kontrollproteine aufgeführt.

**Blau:** His<sub>6</sub>-tag; **grau:** T7-Epitop; **weinrot:** SUMO-tag; **grün:** CBD-tag; gepunktet unterstrichen: Thrombinerkennungs- und Schnittsequenz; **▼:** Position der tag-Abspaltung; **gelb:** gemäß „whole-residue“-Hydrophobizitätsanalyse (MPEx) mit Lipiddoppelmembran interagierender Bereich; doppelt unterstrichen: gemäß „whole-residue“-Hydrophobizitätsanalyse (MPEx)  $\beta$ -Haarnadelstruktur einnehmender Sequenzbereich.

#### 11.1.1 Unveränderte *flank*-Proteine

##### sNT

MGSSHHHHHSSGLVPR▼GSHMGGSSASAAARAAARASA  
GGLGGLGGFGSAAANAAAAANAGAGFGGFGGFGGFGG  
V

Aminosäuren:	75
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	n.b.*
M <sub>w</sub> [kDa]:	6,70
pI:	12,3

##### sCT

MGSSHHHHHSSGLVPR▼GSHMGGGLGGAGPGGLGGLGG  
GLGGLGGGLGGLGGAGGLGGGLGGLGGAGGLGGGLGG  
LGGAGGGAGGAGGAGGAGATAAAAAQAAAAAGAGGNG  
GAAQAAAQAAASAAANSGLGAGAARAAASAAARATVA  
GHGSGTAAAAANAAQ

Aminosäuren:	164
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	n.b.*
M <sub>w</sub> [kDa]:	13,04
pI:	12,3

eNT

MGSSHHHHHHSSGLVPR<sup>▼</sup>GSHMGGFGGIGGIGPGGSSVG  
 GGIGGIGGGVGGVGGIGGVGGPGGIGGIGIGPGFGGG  
 FPGSSASGSGSGSAFGGPGGSSASANAAARANANGG  
 GGFG

Aminosäuren:	115
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	n.b.*
M <sub>w</sub> [kDa]:	9,65
pI:	12,0

eCT

MGSSHHHHHHSSGLVPR<sup>▼</sup>GSHMFPFGFGGFGGFGGASAN  
 AGSSANANAGGPGGGAGAGSSSGANANAGGFPFGGPF  
 GGVGGGPGGAGGPGGVGGGPGGVGGGVGGGPGGVGGG  
 VGGGPGGAGGLGGFGGFGGGSSAGASSSGSASASSGG  
 PFGVLNVGPGGGIGGGSASASAASRA

Aminosäuren:	174
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	n.b.*
M <sub>w</sub> [kDa]:	14,24
pI:	12,0

cwNT

MGSSHHHHHHSSGLVPR<sup>▼</sup>GSHMAAAHAAAVANALGGGG  
 GGSASAAARAAAAARALGGGGGSASAAARAAAAARAL  
 GGGGGASAAARAAAAASALGGGGGFGGLGGLGGGAGG  
 LGGLGGGLGGLGGGVGGGGVGGGVGPGLLGGGGG  
 PGSAGAFGNGNAAAGPGGAGASASAGAFATGGGGFPL  
 PG

Aminosäuren:	187
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	n.b.*
M <sub>w</sub> [kDa]:	15,00
pI:	12,7

cwCT

MGSSHHHHHHSSGLVPR<sup>▼</sup>GSHMGAGEGGAGGAGGGAGG  
 AGGLGGGAGGAGGLGGGFGGLGGGAGGLGGLGGGAGG  
 AGGVGGLGGGVGGLGGVGGLGFGGAGASAGAGANAGA  
 GGAGGSASA

Aminosäuren:	120
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	n.b.*
M <sub>w</sub> [kDa]:	9,22
pI:	7,1

\* n.b.: durch ProtParam rechnerisch nicht bestimmbar, da Tyrosin-, Tryptophan- und Cysteinreste fehlen

### 11.1.2 Optimierte *flank*-Fusionsproteine

#### SUMO-eCT<sub>2Y</sub>

MGSSHHHHHHSSGLVPR<sup>▼</sup>GSHMSDSEVNQEAKPEVKPE  
VKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFK  
RQKEMDSLRLFLYDGIRIQADQTPEDLDMEDNDIIEA  
HREQIGG<sup>▼</sup>SYPGFGGFGGFGGASANAGSSANANAGGPG  
GSAGAGSSSGANANAGGFPGGPFGGVGGGPGGAGGP  
GGVGGGPGGVGGGVGGGPGGVGGVGGGPGGAGGLGG  
FGGFGGSSAGASSSGSASASSGGPFGLNVGPGGGI  
GGGSASASAASRAYA

Aminosäuren:	274
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	4470
M <sub>w</sub> [kDa]:	25,69
pI:	6,12

#### SUMO-cwNT<sub>2Y</sub>

MGSSHHHHHHSSGLVPR<sup>▼</sup>GSHMSDSEVNQEAKPEVKPE  
VKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFK  
RQKEMDSLRLFLYDGIRIQADQTPEDLDMEDNDIIEA  
HREQIGG<sup>▼</sup>SYAAHAAAVANALGGGGGSASAAARAAA  
AARALGGGGGSASAAARAAAAAARALGGGGGSASAAARA  
AAAASALGGGGFGGLGGLGGGAGGLGGLGGLGGLG  
GGGVGGGVGGGVGPGGLLGGGGPGSAGAFGNGNA  
AAGPGGAGASASAGAFATGGGGFPLPGYA

Aminosäuren:	288
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	4470
M <sub>w</sub> [kDa]:	26,60
pI:	7,2

#### SUMO-cwCT<sub>2Y</sub>

MGSSHHHHHHSSGLVPR<sup>▼</sup>GSHMSDSEVNQEAKPEVKPE  
VKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFK  
RQKEMDSLRLFLYDGIRIQADQTPEDLDMEDNDIIEA  
HREQIGG<sup>▼</sup>SYGAGEGGAGGAGGGAGGAGGLGGGAGGAG  
GLGGFGGLGGGAGGLGGLGGGAGGAGGVGGLGGGVG  
GLGGVGGGLGFGGAGASAGAGANAGAGGAGGSASAYA

Aminosäuren:	221
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	4470
M <sub>w</sub> [kDa]:	20,81
pI:	5,8

#### SUMO-eCT(F106W)

MGSSHHHHHHSSGLVPR<sup>▼</sup>GSHMSDSEVNQEAKPEVKPE  
VKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFK  
RQKEMDSLRLFLYDGIRIQADQTPEDLDMEDNDIIEA  
HREQIGG<sup>▼</sup>SPGFGGFGGFGGASANAGSSANANAGGPGG  
SAGAGSSSGANANAGGFPGGPFGGVGGGPGGAGGPG  
GVGGGPGGVGGGVGGGPGGVGGVGGGPGGAGGLGGF  
GGWGGSSAGASSSGSASASSGGPFGLNVGPGGGIG  
GGSASASAASRA

Aminosäuren:	271
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	6990
M <sub>w</sub> [kDa]:	25,33
pI:	6,1

## SUMO-cwCT(F34W)

MGSSHHHHHHSSGLVPR<sup>▼</sup>GSHMSDSEVNQEAKPEVKPE  
 VKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAK  
 RQKEMDSLRFlyDGIRIQADQTPEDLDMEDNDIIEA  
 HREQIGG<sup>▼</sup>SMGAGEGGAGGAGGGAGGAGGLGGGAGGAG  
 GLGGGWGGLGGGAGGLGGLGGGAGGAGGVGGLGGGVG  
 GLGGVGGLGFGGAGASAGAGANAGAGGAGGSASA

Aminosäuren:	219
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	6990
M <sub>w</sub> [kDa]:	20,59
pI:	5,8

## SUMO-cwCT(F34W)-CBD

MGSSHHHHHHSSGLVPR<sup>▼</sup>GSHMSDSEVNQEAKPEVKPE  
 VKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAK  
 RQKEMDSLRFlyDGIRIQADQTPEDLDMEDNDIIEA  
 HREQIGG<sup>▼</sup>SMGAGEGGAGGAGGGAGGAGGLGGGAGGAG  
 GLGGGWGGLGGGAGGLGGLGGGAGGAGGVGGLGGGVG  
 GLGGVGGLGFGGAGASAGAGANAGAGGAGGSASA<sup>▼</sup>CIT  
 GDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDR  
 HGNPVLADRLFHSGEHPVYTVRTVEGLRVGTANHPL  
 LCLVDVAGVPTLLWKLIDEIKPGDYAVIQSAFSVDC  
 AGFARGKPEFAPTTYTVGVPLVRFLEAHHRDPDAQA  
 IADELTDGRFYYAKVASVTDAGVQPVYSLRVDADHA  
 FITNGFVSHATGLTGLNSGLTTNPGVSAWQVNTAYTA  
 GQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQLQ

Aminosäuren:	479
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	Cys <sub>red</sub> 42400 Cys <sub>ox</sub> 42650
M <sub>w</sub> [kDa]:	48,43
pI:	5,7

## 11.1.3 Fusions-tags

SUMO-tag (inkl. His<sub>6</sub>-tag)

MGSSHHHHHHSSGLVPR<sup>▼</sup>GSHMSDSEVNQEAKPEVKPE  
 VKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAK  
 RQKEMDSLRFlyDGIRIQADQTPEDLDMEDNDIIEA  
 HREQIGG

Aminosäuren:	118
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	1490
M <sub>w</sub> [kDa]:	13,43
pI:	6,0



**CBD-taq (inkl. *Mxe* *gyrA* Inteinsequenz)**

CITGDALVALPEGESVRIADIVPGARPNSDNAIDLKV  
 LDRHGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTAN  
 HPLLCLVDVAGVPTLLWKLIDEIKPGDYAVIQRSAFS  
 VDCAGFARGKPEFAPTTYTVGVPLVRFLEAHRDPD  
 AQAIADELTDGRFYAKVASVTDAGVQPVYSLRVDTA  
 DHAFITNGFVSHATGLTGLNSGLTTNPGVSAWQVNTA  
 YTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQL  
 Q

Aminosäuren:	260
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	Cys <sub>red</sub> 35410 Cys <sub>ox</sub> 35660
M <sub>w</sub> [kDa]:	27,86
pI:	5,7

**11.1.4 Verwendete Kontrollproteine****11.1.4.1 Spinnenseidenproteinanaloge****eADF4(C16)**

MASMTGGQQMGRGSMGSSAAAAAASGPGGYGPENQ  
 GPSGPGGYGPGGPGSSAAAAAASGPGGYGPENQGP  
 SGPGGYGPGGPGSSAAAAAASGPGGYGPENQGPSG  
 PGGYGPGGPGSSAAAAAASGPGGYGPENQGPSGPG  
 GYGPGGPGSSAAAAAASGPGGYGPENQGPSGPGGY  
 GPGGPGSSAAAAAASGPGGYGPENQGPSGPGGYGP  
 GPGGPGSSAAAAAASGPGGYGPENQGPSGPGGYGPGG  
 PGSSAAAAAASGPGGYGPENQGPSGPGGYGPGGPG  
 SSAAAAAASGPGGYGPENQGPSGPGGYGPGGPGSS  
 AAAAAAASGPGGYGPENQGPSGPGGYGPGGPGSSAA  
 AAAAAASGPGGYGPENQGPSGPGGYGPGGPGSSAAAA  
 AAAASGPGGYGPENQGPSGPGGYGPGGPGSSAAAAA  
 AASGPGGYGPENQGPSGPGGYGPGGPGSSAAAAA  
 SGPGGYGPENQGPSGPGGYGPGGPGSSAAAAAASG  
 PGGYGPEHQGPSGPGGYGPGGPGSSAAAAAASGPG  
 GYGPEHQGPSGPGGYGPGGPG

Aminosäuren:	576
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	47680
M <sub>w</sub> [kDa]:	47,70
pI:	3,5

**Sfl**

MASMTGGQQMGRGSMGPGGAGPGGYGPGGSGPGGYGP  
 GSGPGGYGPGGAGPGGYGPGGSGPGGYGPGGSGPGG  
 YGPGGAGPGGYGPGGSGPGGYGPGGSGPGGYGPGGAG  
 PGGYGPGGSGPGGYGPGGSGPGGYGPGGAGPGGYGPG  
 GSGPGGYGPGGSGPGGYGPGGAGPGGYGPGGSGPGGY  
 GPGGSGPGGYGAGGAGGAGGSGGAGGSGGAGGAGGA  
 GSGGAGGSGGTTIIEDLDITIDGADGPITISEELTI  
 GPGGAGGPYPGGAGGPYPGGAGGPYPGGAGGPYPG  
 PGGAGGPYPGGAGGPYPGGAGPGGYGPGGSGPGGY  
 GPGGSGPGGYGPGGAGPGGYGPGGSGPGGYGPGGSGP  
 GGY

Aminosäuren:	373
$\epsilon_{280\text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	44700
M <sub>w</sub> [kDa]:	29,71
pI:	3,6

## 11.1.4.2 Weitere Proteine

BSA\*

DTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQCPFD  
 EHVKLVNELTEFAKTCVADESHAGCEKSLHTLFGDEL  
 CKVASLRETYGDMADCCEKQEPERNECFLSHKDDSPD  
 LPKLKPDNTLCDEFKADEKKFWGKYLYEIAARRHPYF  
 YAPELLYYANKYNGVFQECCQAEDKGACLLPKIETMR  
 EKVLAASSARQRLRCASIQKFGERALKAWSVARLSQKF  
 PKAEFVEVTKLVTDLTKVHKECCHGDLLECADDRADL  
 AKYICDNQDTISSKLKECCDKPILLEKSHCIAEVEKDA  
 IPENLPPLTADFAEDKDVCKNYQEAKDAFLGSFLYEY  
 SRRHPEYAVSVLLRLAKEYEATLEECCA KDDPHACYS  
 TVFDKLKHLVDEPQNLIKQNC DQFEKLGEYGFQNALI  
 VRYTRKVPQVSTPTLVEVSRSLGKVGTRCCTKPESER  
 MPCTEDYLSLILNRLC V LHEKTPVSEKVT KCCTESLV  
 NRRPCFSALTPDETYVPKAFDEKLFTFHADI CTLPDT  
 EKQIKKQTALVELLKHKPKATEEQLKTVMENFVAFVD  
 KCCAADDKEACFAVEGPKLVVSTQTALA

\*Quelle: UniProtKB, P02769[25-607]

Aminosäuren:	583
$\epsilon_{280 \text{ nm}}$ [M <sup>-1</sup> cm <sup>-1</sup> ]:	Cys <sub>red</sub> 40800 Cys <sub>ox</sub> 42925
M <sub>w</sub> [kDa]:	66,43
pI:	5,6

## 12 Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbstständig unter ausschließlicher Nutzung der darin angegebenen Quellen und Hilfsmittel angefertigt habe. Des Weiteren erkläre ich, dass ich weder diese noch eine gleichartige Dissertation jemals an einer Hochschule eingereicht und mich noch nie einer Doktorprüfung unterzogen habe. Ich erkläre darüber hinaus, dass ich weder die Hilfe eines gewerblichen Promotionsberaters in Anspruch genommen habe noch dieses in Zukunft tun werde.

Markus Heim

Bayreuth, Februar 2014